A STUDY OF INTERACTIONS BETWEEN NEAR-INERTIAL INTERNAL WAVES AND MESOSCALE-TO-SUBMESOSCALE FLOWS

A Dissertation

by

ZHAO JING

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Ping Chang
Co-Chair of Committee,	Lixin Wu
Committee Members,	Ramalingam Saravanan
	Steven DiMarco
Head of Department,	Debbie Thomas

May 2016

Major Subject: Oceanography

ABSTRACT

In this dissertation, we explore interactions between near-inertial internal waves (NIWs) and mesoscale-to-submesoscale flows. Three different topics under this broad subject are investigated using theoretical, numerical and observational approaches.

The downward radiation of NIWs from the mixed layer to deep ocean in an idealized baroclinic geostrophic flow is theoretically analyzed based on the Young-Ben Jelloul (YBJ) equation. It is found that the dispersion of NIWs in the presence of baroclinic flow is achieved mainly through the phase separation among different horizontal and vertical modes. Both the eigen-frequency differences and mode-mode interferences contribute to the phase separation with the interferences locally in the modal space playing a much more dominant role than the nonlocal mode-mode interferences.

Data from long-term mooring array and high-resolution numerical simulations in the Gulf of Mexico are used to analyze energy exchange between near-inertial internal waves and mesoscale eddies. Both the observations and numerical simulations reveal a permanent energy transfer from mesoscale eddies to NIWs below the mixed layer. In particular, this permanent energy transfer mainly occurs when the Okubo-Weiss (OW) parameter is positive. Further analysis suggests that the wave capture mechanism plays a key role in interactions between NIWs and mesoscale eddies. NIWs become highly anisotropic when the OW parameter is positive. The observed probability density

function of propagation direction of NIWs is consistent with the predictions from the wave capture theory.

Submesoscale O(<10 km) motions and their interactions with NIWs are studied theoretically and numerically using high-resolution numerical simulations. Submesoscale fronts (SMFs) with energetic vertical motions in the ocean interior are found to be closely associated with the NIWs. A dynamic mechanism for the SMF development in the presence of background NIWs is proposed. It shows that in convergence (downwelling) regions of NIWs, energy flux of the submesoscale motions converges and the energy is transferred from the NIWs to submesoscale motions, leading to enhanced submesoscale vertical velocity. The opposite is true in divergence (upwelling) zones of NIWs. The underlying dynamics can be understood in terms of wave action conservation of submesoscale motions in the presence of background NIWs.

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to my advisor, Dr. Ping Chang. Without his support and encouragement, this dissertation would not have been possible. I was incredibly lucky to have him as an advisor for his broad knowledge, perceptive insight into scientific research and responsible scientific attitude. I would also like to thank my Co-Chair Dr. Lixin Wu as well as my committee members Dr. Ramalingam Saravanan and Dr. Steven DiMarco for their constructive comments and tremendous knowledge.

I would like to thank my colleagues during my Ph.D. study over the past four years. A special thank goes to Dr. Xiaohui Ma who instructed me to run the Regional Ocean Modelling System. Without her help, I would not have completed the numerical simulations used in this dissertation. I also want to thank Raffaele Montuoro who helped me configure the model simulations in the Gulf of Mexico. I would like to extend my thanks to other colleagues in my research group, my friends and the faculty in the Department of Oceanography for their aids during my graduate study.

I am grateful to China Scholarship Council for their financial support. I am also grateful to the Texas A&M Supercomputing Facility and the Texas Advanced Computing Center (TACC) at the University of Texas at Austin, which provided the computing resources used in this dissertation.

Finally, I would like to express my heartfelt gratitude to my family for their solid support and unconditional love.

NOMENCLATURE

ADCP	Acoustic Doppler Current Profiler
CTD	Conductivity-Temperature-Depth
f	Corilois
GISR	Gulf Integrated Spill Research Consortium
GoM	Gulf of Mexico
hycom	Hybrid Coordinate Ocean Model
IP	Inertial Period
K1	Luni-Solar Diurnal
KESS	Kuroshio Extension System Study
M2	Principal Lunar
N	Buoyancy Frequency
NIW	Near-Inertial Internal Wave
O1	Principal Lunar Diurnal
OW	Okubo-Weiss Parameter
PSI	Parametric Subharmonic Instability
PWP	Price-Weller-Pinkel
RDI	Rowe-Deines Instruments
RMS	Root Mean Square
ROMS	Regional Ocean Modeling System
S2	Principal Solar
SMF	Submesoscale Front

SST	Sea Surface Temperature
TW	Terawatt
YJB	Young and Ben Jelloul
β	Planetary Vorticity Gradient

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION

1.1 Background

1.1.1 Near-inertial internal waves

Gravity waves on the sea surface are a phenomenon of common experience. Distortions of an otherwise quiescent sea surface feel the restoring force of gravity and disperse in forms of travelling surface waves. A stably stratified ocean can similarly support gravity waves in its interior, which are referred to as internal gravity waves or internal waves for short. They have a larger amplitude compared to surface gravity waves due to reduced restoring forces, capable of displacing isopycnals in the ocean interior by tens or hundreds of meters (Garret and Munk 1979).

Frequencies of internal waves range from the Coriolis frequency f (also referred to as the inertial frequency) to the buoyancy frequency N. Near f the fluid motion is close to horizontal, with the Coriolis force dominating. In contrast, near N the motion is close to vertical, with buoyancy providing the main restoring force. Observations of ocean currents usually exhibit a continuous frequency spectrum between f - N with pronounced spectral peaks around f and frequencies of major tidal constituents (Figure 1.1). The peak around the inertial frequency corresponds to near-inertial internal waves (NIWs) while the peaks at the major tidal constituents are formed by internal tides.

In the upper ocean, NIWs dominate the internal wave filed, containing half of the kinetic energy and a substantial portion of shear variance in the internal wave field

(Alford and Whitmont 2007; Ferrari and Wunsch 2009). They appear nearly everywhere in the ocean with their energy level rising and falling on short time scales and varying strongly with location (Fu 1981; Alford and Whitmont 2007). Global map of near-inertial kinetic energy derived from observations of surface drifters reveals energetic NIWs in the Gulf Stream, Kuroshio extension region and Southern Ocean (Chaigneau et al. 2008). A two-year time series of near-inertial kinetic energy at the ocean station Papa exhibits an evident seasonal cycle as well as numerous strong-NIW events lasting 5-20 inertial periods (Alford et al. 2012).

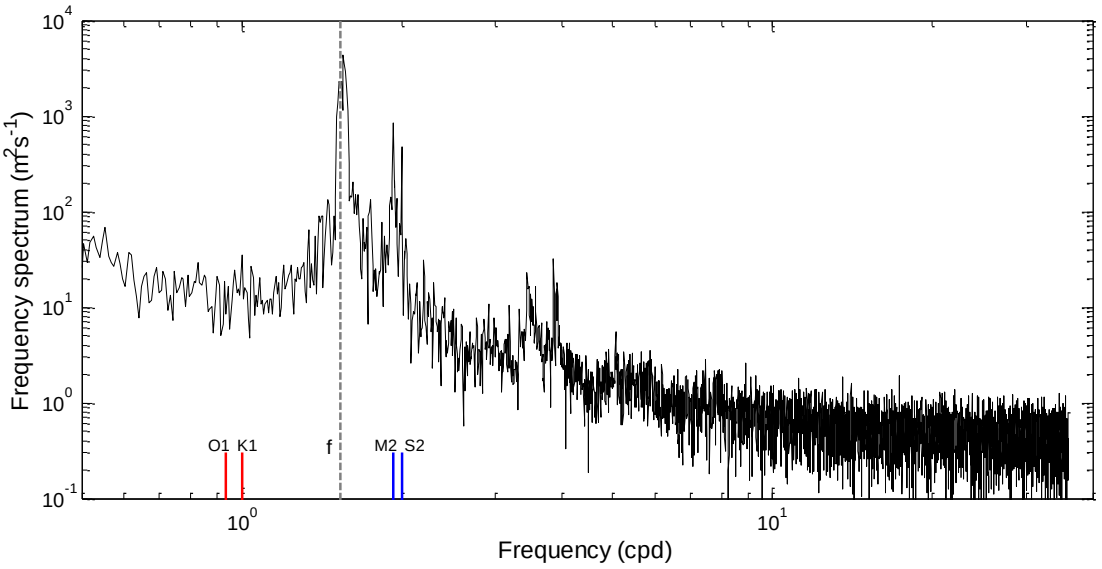

Figure 1.1 Frequency spectrum of surface ocean current at the ocean station Papa (50°N, 145°W). The inertial, diurnal and semi-diurnal tidal periods are marked.

NIWs are generated by a variety of mechanisms, including wind forcing (D'Asaro 1985; Alford 2003a), parametric subharmonic instability (PSI) at critical latitudes (Alford 2008; MacKinnon and Winters 2005; MacKinnon et al. 2013), lee-wave

formation by geostrophic flows over bottom topography (Nikurashin and Ferrari 2010a and 2010b), and spontaneous generation of unbalanced motions (Vanneste 2013; Alford et al. 2013; Nagai et al. 2015). Among these mechanisms, winds provide a major energy source for NIWs. As a natural resonant frequency of fluids on a rotating planet, wind stress fluctuations with frequencies in near-inertial band can resonantly force NIWs in surface mixed layer. Global estimates for the wind work on NIWs range from 0.3-1.4 TW (Watanabe and Hibiya 2002; Alford 2003a; Jiang et al. 2005; Simmons and Alford 2012; Furuichi et al. 2008; Rimac et al. 2013). Much of the spread in the estimated wind work can be attributed to differences in the wind products used in these estimates. Nevertheless, all the estimates reveal that regions of strong wind work and near-inertial kinetic energy in surface mixed layer are collocated with each other (Chaigneau et al. 2008), suggesting the important role of winds in generating NIWs.

Most of wind work on NIWs occurs when fast moving weather systems (e.g., tropical cyclones, fronts and midlatitude storms) pass over the ocean. The oceanic response may be conveniently divided into two stages due to the relatively long time scale of baroclinic geostrophic adjustment compared to the wind forcing time scale (Price et al. 1994). The first stage is considered as a "forced stage" when weather systems pass over and near-inertial energy is rapidly deposited to surface mixed layer. The second stage is regarded as a "relaxation stage" when wind-input near-inertial energy disperses in the form of NIWs. The latitudal dependence of f, i.e., the β effect, plays a key role in horizontal propagation of NIWs (D'Asaro 1989). According to ray tracing theory, the horizontal group velocity of NIWs increases as they propagate

equatorwards due to the decreased f. In contrast, poleward-propagating NIWs are associated with a decreasing horizontal group velocity and are reflected at a turning latitude where f is equal to the wave frequency. Due to the asymmetry between the equatorward and poleward propagation, propagation of NIWs is nearly always equatorward in the ocean (Garret 2001).

At the relaxation stage, the wind-generated near-inertial energy radiates downwards from mixed layer to thermocline and deep ocean. Gill (1984) developed a modal formalism for describing the downward radiation of near-inertial energy. Essentially, near-inertial current is projected onto a spectrum of vertical modes. At the beginning of relaxation stage, phases of different vertical modes work constructively so that all the near-inertial kinetic energy is confined to the surface mixed layer. Then different vertical modes dephase due to their differentiated wave frequencies, leading to downward radiation of near-inertial energy into ocean interior. In the real ocean, the dephasing time scale ranges from a few days to several weeks and is affected by the horizontal scale of wind forcing (Gill 1984), the β effect (D'Asaro 1989) as well as the vorticity of background flow (e.g., Balmforth et al. 1998).

The modal partition of near-inertial kinetic energy onto vertical modes depends on the stratification and mixed layer depth. In the real ocean, near-inertial kinetic energy projects predominantly onto the lowest modes (Gill 1984) although high modes with a vertical wavelength of less than several hundred meters make a more important contribution to near-inertial shear variance (Alford and Gregg 2001; Jing and Wu 2014). NIWs of high modes typically dissipate locally due to the strong vertical shear that

generates elevated turbulence. In contrast, NIWs of low modes can travel a long distance from their source (Alford 2003b). It remains poorly understood where and through what mechanism the near-inertial kinetic energy carried by low modes dissipates. Possible candidates include nonlinear wave-wave interactions and reflection at critical topography.

1.1.2 Interactions between near-inertial internal waves and mesoscale flows

Storm tracks are remarkably coincident with regions of strong mesoscale eddy activities (Zhai et al. 2005). In addition, NIWs are associated with a slow horizontal and vertical group velocity due to the small difference between their wave frequencies and *f*. Therefore, there are ample opportunities for NIWs to interact with mesoscale flows. Previous studies suggested that vorticity of background flow has a significant influence on NIW propagation. In the presence of background vorticity, the lower bound of the internal wave frequency band is shifted from f to an effective Coriolis frequency $f_{eff} = f + \varsigma/2$ (Kunze 1985) where ς is the vorticity of background flow. According to ray tracing theory, the horizontal heterogeneity of f_{eff} drives NIWs to propagate from positive-ς regions to negative-ς regions in the northern hemisphere and the opposite is true in the southern hemisphere. This leads to concentration of near-inertial energy in anticyclonic regions (Balmforth et al. 1998; Klein and Treguier 1995; Klein and Smith 2001; Klein et al. 2004). In particular, NIWs generated in an anticyclonic eddy can be trapped within it when the wave frequency is lower than the surrounding f_{eff} (Kunze 1985). These theoretical arguments have been confirmed by observations (Elipot et al.

2010; Rainville and Pinkel 2004; Jing and Wu 2014) and numerical simulations (Lee and Niiler 1998; Zhai et al. 2005; Zhai et al. 2007).

Background flow vorticity also plays an important role in vertical dispersion of NIWs. Numerical simulations suggest that radiation of near-inertial energy from mixed layer into deep ocean is strongly enhanced by the presence of mesoscale eddies, with anticyclonic mesoscale eddies acting as a conduit to the deep ocean (Lee and Niiler 1998; Zhai et al. 2005). Mooring observations in the Kuroshio extension region reveal that the downward near-inertial energy flux into the deep ocean accounts for 45%-62% of the wind work on NIWs (Jing and Wu et al. 2014). In contrast, less than 33% of wind-input near-inertial energy is able to radiate into the deep ocean based on the observations collected in the Northeast Pacific where mesoscale eddy activities are much weaker (Alford et al. 2012), confirming the important role of mesoscale eddies in promoting the downward radiation of near-inertial energy.

Mesoscale eddies do not only affect horizontal and vertical dispersion of NIWs but are also their potential energy sources. Global estimates for the wind work on geostrophic flows are about 1 TW (Wunsch and Ferrari 2004). It remains unclear how the energy is eventually dissipated. On one hand, energy transfer from mesoscale eddies to NIWs could be a significant sink of energetic mesoscale eddies (Ferrari and Wunsch 2009; Polzin 2010). On the other hand, it may provide an important energy source for NIWs which may eventually sustain diapycnal mixing. Observations indicated a permanent energy transfer from mesoscale eddies to NIWs with great uncertainties in magnitude (Frankignoul 1976; Ruddick and Joyce 1979; Frankignoul and Joyce 1979;

Polzin 2010). Several mechanisms have been proposed for the permanent energy exchange between NIWs and mesoscale eddies, including wave breaking near critical layers (Bretherton 1966), relaxation effects through nonlinear wave-wave interactions (Müller 1976), conservation of wave angular momentum (Weller 1982), wave capture (Bühler and McIntyre 2005), and simultaneous generation (Vanneste 2013).

1.1.3 Role of near-inertial internal waves in the global climate system

A key role of NIWs in the global climate system is their contribution to diapycnal mixing which affects uptake of heat and carbon by the oceans, as well as oceanic response to climate change (Gregory 2000; Huang 1999; Jayne 2009; Richards et al. 2009; Saenko and Merrifield 2005; Sarmiento and Toggweiler 1984; Wunsch and Ferrari 2004). NIWs generate strong vertical shear at the base of mixed layer. This enhances the mixing through shear instability, which is a major mechanism for mixed layer deepening and cooling during strong wind forcing (Greatbatch, 1984; Price et al., 1986; Plueddemann and Farrar 2006). In a recent numerical study by Jochum et al. (2013), it was found that NIWs deepen the mixed layer depth by up to 30%. This further affects the sea surface temperature (SST) and could impact the global climate system through air-sea coupling.

NIWs may also play an important role in furnishing diapycnal mixing in the thermocline and deep ocean (Wunsch and Ferrari 2004). According to a 1-D advection-diffusion model study, Munk and Wunsch (1998) estimated that about 2 TW (1TW= 10^{12} W) energy is required for sustaining diapycnal mixing below the thermocline. The total tidal dissipation rate in the open ocean is, however, estimated to be only about 1

TW (Egbert and Ray 2001; Jayne and St. Laurent 2001). NIWs may make important contribution to the remaining 1TW energy. As mentioned in Section 1.1.1, global estimates for wind work on NIWs range from 0.3-1.4 TW, which is comparable to the value of tidal dissipation. If a substantial portion of wind-input near-inertial energy was able to radiate into the deep ocean, NIWs would play an important role in furnishing abyssal diapycnal mixing. This conjecture has been, to some extent, confirmed by direct and indirect observational evidence. Global maps of ocean mixing inferred from Argo profiles using a finescale parameterization for mixing by internal waves reveal that the subthermocline diapycnal mixing is enhanced in the western boundary current regions associated with strong near-inertial wind work (Whalen et al. 2012). In particular, the diapycnal mixing there exhibits a pronounced seasonal cycle, which is in phase with the seasonal variation of near-inertial wind work (Jing and Wu 2010; Whalen et al. 2012). Furthermore, in a recent study by Jing and Wu (2014), a tight correlation of diapycnal diffusivity and near-inertial shear variance was found in the upper 1500 m of Kuroshio extension region. All the evidence suggests that the NIWs probably make an important contribution to diapycnal mixing in the deep ocean.

1.2 Thesis Objectives

The focus of this study is on the interactions between NIWs and mesoscale-to submesoscale flows. As discussed in Section 1.1.2, propagation of NIWs is significantly affected by mesoscale eddies. But the underlying mechanisms are still not well understood. In particular, most of existing theoretical studies are based on a simple

assumption that mesoscale eddies are barotropic (Balmforth et al. 1998; Klein and Treguier1995; Klein and Smith 2001; Klein et al. 2004). In reality, mesoscale eddies are baroclinic rather than barotropic. It remains unclear whether vertical structure of mesoscale eddies can have a significant influence on the radiation of near-inertial energy. In particular, NIWs of different vertical modes are potentially able to interfere with each other in the presence of a baroclinic background flow, but evolve independently otherwise. This brings us to our first research question:

1) What is the difference between horizontal and vertical dispersion of NIWs in the presence of a barotropic and baroclinic geostrophic flow. Do interferences among NIWs of different vertical modes play an important role in the radiation of near-inertial energy in the presence of a baroclinic geostrophic flow?

Next, we investigate energy exchange between mesoscale eddies and NIWs. Many mechanisms have been proposed for permanent energy transfer from mesoscale eddies to NIWs. But so far it remains uncertain which mechanism plays a dominant role due to the limited observations. For the observations, difficulties in analyzing the energy exchange between mesoscale eddies and NIWs result from the evaluation of horizontal gradients of mesoscale eddies. During 2012-2013, a 6-mooring array was deployed in the northern Gulf of Mexico. Each mooring was equipped with an Acoustic Doopler Current Profiler (ADCP) to measure ocean currents in subthermocline. The distance between neighboring moorings is ~25 km, which is able to resolve bulk of the strain and vorticity variance associated with mesoscale eddies. This provides a unique opportunity to study our second research question:

2) What is the dominant mechanism responsible for permanent energy exchange between mesoscale eddies and NIWs and how important this energy exchange is to kinetic energy budget of NIWs?

Remarkably little is known for the interactions between NIWs and submesoscale (<10 km) flows. But such interactions could be important to the vertical transport processes in the ocean as the submesoscale flows are characterized by a much smaller horizontal/vertical aspect ratio than NIWs and thus associated with stronger vertical velocity. In a recent modelling study in the Gulf of Mexico, pronounced submesoscale vertical motions in the ocean interior are found to be closely associated with energetic NIWs (Zhong and Bracco 2013). This close association may imply a dynamic linkage between them. However, underlying mechanisms responsible for this linkage remain unexplored. This motivates our third research question:

3) Can vertical motions of submesoscale flows be modulated by NIWs? If so, what is the modulation mechanism?

1.3 Outline of the Thesis

This thesis comprises five chapters and five appendixes. In Chapter II, we address the first research question. Dispersion of NIWs in idealized baroclinic and barotropic geostrophic flows is theoretically analyzed based on the equation proposed by Young and Ben Jelloul (1997) (the YJB equation). We explore the mechanisms responsible for the dispersion of NIWs in baroclinic geostrophic flows and compare the differences between baroclinic and barotropic flows.

In Chapter III, we address the second research question. Energy exchange between NIWs and mesoscale eddies is computed based on both observations and high-resolution numerical simulations in the Gulf of Mexico. The relative importance of various mechanisms discussed in Section 1.1.2 is evaluated.

In Chapter IV, we address the third research question. We first design an idealized high-resolution numerical simulation to illustrate the relationship between submesoscale fronts (SMFs) and NIWs. Theoretical solutions governing SMF – NIW interactions are then developed. The validity of the theoretical solutions is tested using both the idealized numerical simulation and realistic numerical simulation in the Gulf of Mexico. This chapter has been submitted to Journal of Physical Oceanography.

In Chapter V, we summarize the most important findings of the study and discuss the questions remaining to be answered in future studies

CHAPTER II

DISPERSION OF NEAR-INERTIAL INTERNAL WAVES IN BAROCLINIC GEOSTROPHIC FLOWS

2.1 Introduction

NIWs are excited and sustained primarily by the wind work on near-inertial motions in the mixed layer (Wunsch and Ferrari 2004). The subsequent downward and horizontal radiation of NIWs enhances shear in ocean interior and triggers mixing events locally. Distinct seasonality of diapycnal mixing has been observed in the upper ocean and is significantly correlated to the near-inertial wind work (Jing and Wu 2010 and 2014). Understanding the dispersion of NIWs is essential to parameterize the diapycnal mixing in the thermocline and deep ocean, which may further improve models' representation and simulation of large-scale ocean circulation and its response to climate variability/change (Saenko and Merrifield 2005; Wunsch and Ferrari 2004).

Most of the near-inertial wind work occurs when fast moving wind storms input strong momentum into the ocean surface mixed layer. The response of the ocean may be conveniently divided into two stages due to the relatively long time scale of the baroclinic geostrophic adjustment process compared to that of wind forcing (Price 1983). The first stage is considered as "forced stage" when a storm is passing overhead and the wind energy is rapidly deposited into the mixed layer. The second stage is considered as "relaxation stage" when the energy in the mixed layer radiates downwards and outwards in forms of NIWs. The forced stage determines the near-inertial energy input by winds

and is relatively well understood. The theoretical models such as the slab mixed layer model (Pollard and Millard 1970; Alford 2003) and the Price-Weller-Pinkel (PWP) model (Price et al. 1986) have shown good skills in simulating the near-inertial response in the mixed layer to surface wind forcing (D'Asaro et al. 1985; Price et al. 1986; Plueddemann and Farrar 2006; Alford et al. 2012).

The relaxation stage, characterizing dispersion of NIWs, has not been fully understood but has significant influences on the potential role of NIWs in diapycnal mixing in the ocean interior. Gill (1984) examined the dynamics during the relaxation stage by linearly projecting slab mixed layer velocity onto vertical modes. In this linear model framework, there is no background flow and β effect. Downward radiation of near-inertial energy is achieved by phase separation among different vertical modes due to the difference in the associated eigen-frequencies $\omega_n = f + c_n^2 k_H^2 /(2f)$ where k_H denotes horizontal wavenumber of NIWs and c_n represents the gravity speed of the *n*-th vertical mode. At the beginning of relaxation stage, phases of different vertical modes work constructively so that all the near-inertial kinetic energy is confined to the surface mixed layer (Figure 2.1a). Then different vertical modes dephase due to their differentiated eigen-frequencies. In particular, as c_1 is typically much larger than c_n associated with higher vertical modes, the first vertical mode becomes approximately 90^o out of phase with the higher modes after a period of $f\pi/(c_1^2 k_h^2)$. At this stage, near-inertial current emerges in the thermocline and deep ocean while the mixed layer near-inertial current amplitude is reduced (Figure 2.1b), corresponding to a downward

radiation of near-inertial energy from the mixed layer into the thermocline and deep ocean.

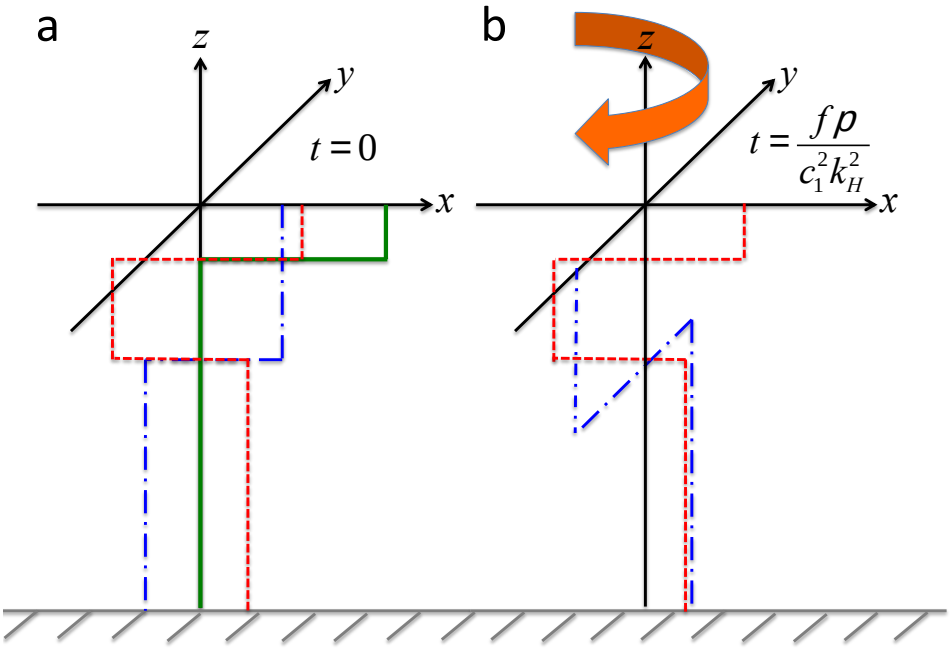

Figure 2.1 A schematic illustrating the downward radiation of near-inertial energy from the mixed layer into the thermocline and deep ocean through the phase separation among vertical modes. The blue and red lines represent the vertical profiles of near-inertial current associated with the first vertical mode and with the higher vertical modes (a) at t=0 (the beginning of relaxation stage) and (b) at $t = f\pi/(c_1^2 k_h^2)$. The green line in (a) denotes the vertical profile of near-inertial current at t=0.

The above discussion suggests that the dephasing time scale is proportional to the squared wavelength of NIWs. For NIWs with a horizontal scale of O(100 km), Gill's solution predicts a time scale of several days for the decay of mixed-layer near-inertial energy through downward radiation, consistent with observations (Pollard and Millard 1970; Price 1983). However, as horizontal scale of midlatitude synoptic storms is typically of O(1000 km), the predicted dephasing time scale should be of O(100-1000

day). Such a long time scale implies that much of the near-inertial energy should be dissipated within mixed layer or near mixed layer base without significant downward energy propagation, contradicting the available *in situ* observations (D'Asaro et al. 1995; Alford et al. 2012; Silverthorne and Toole 2009).

The deficiency of the Gill's solution may be due to the negligence of the β effect or/and background flow. The influence of β effect can be easily understood based on ray tracing approach that gives wave energy propagation path (Lighthill 1978). When NIWs propagate towards the equator, the meridional wavenumber increases linearly with time due to the β effect, accelerating the phase separation among vertical modes (D'Asaro 1989). The β effect imposes an upper limit of several weeks for the decay of near-inertial energy in the mixed layer, supported by the Ocean Storms Experiment (D'Asaro et al. 1995).

Another factor affecting NIW propagation is the vorticity of background geostrophic flows. Ray tracing studies suggest that background geostrophic vorticity ς can shift the lower bound of the internal waveband from the planetary value of the Coriolis frequency f to an effective Coriolis frequency $f_{eff} = f + \varsigma/2$ (Kunze 1985). Energetic NIWs have been shown to concentrate in the negative-ς region[1] according to numerical model simulations (Lee and Niiler 1998, Zhai et al. 2005 and 2007) and theoretical studies (Klein and Treguier 1995; Balmforth et al. 1998). Observations (Kunze 1995; Jing and Wu 2013) indicated that diapycnal mixing in the upper ocean

[1] The analysis in this Chapter is confined to the Northern Hemisphere.

becomes more energetic within anticyclonic eddies, further supporting the trapping of near-inertial wave energy in negative-ς regions.

Background geostrophic flows also play an important role in vertical propagation of NIWs and may work more effectively than the β effect in regions with strong mesoscale eddies. Numerical simulations (Zhai et al. 2005 and 2007) indicate enhanced downward radiation of near-inertial energy in the presence of energetic mesoscale eddies. Balmforth et al. (1998) studied dispersion of NIWs in an idealized barotropic geostrophic flow and found that the decay of mixed-layer near-inertial energy is significantly accelerated. The acceleration is mainly due to the modification of eigen-frequencies of NIW modes by geostrophic vorticity. Small-scale geostrophic vorticity impresses its horizontal structure on NIWs, amplifying eigen-frequency difference among modes and further reducing the dephasing time scale.

In the reality, the mesoscale geostrophic motion is baroclinic rather than barotropic. It remains unclear whether the vertical structure of geostrophic flows has a significant influence on the dispersion of NIWs in the mixed layer and deep ocean. This problem is valuable in practice as geostrophic flows derived from satellites are only available at the sea surface. In this chapter, we advance previous theoretical studies (e.g., Klein and Treguier 1995; Balmforth et al. 1998) by analyzing the dispersion of NIWs in a background baroclinic geostrophic flow. We will show that NIWs of different vertical modes evolve independently in a barotropic geostrophic flow, but are able to interfere with each other in a baroclinic geostrophic flow. In this case, phase separation among different modes of NIWs due to eigen-frequency differences may not be the only

mechanism responsible for redistribution of near-inertial wave energy. To facilitate the theoretical analyses, an idealized baroclinic geostrophic flow is used. The geostrophic flow is identical to that used by Balmforth et al. (1998) except that it is baroclinic rather than barotropic. The chapter is organized as follows. In section 2.2, we briefly introduce the idealized baroclinic flow and the analytical tools. Section 2.3 presents the dispersion of near-inertial energy initially confined to surface mixed layer. Mechanisms responsible for the dispersion of NIWs are explored in Section 2.4. Discussion is presented in Section 2.5 followed by a summary in Section 2.6.

2.2 Methodology

2.2.1 *The YBJ equation*

Young and Ben Jelloul (1997) proposed an equation (known as the YBJ equation) describing evolution of NIWs in geostrophic flows. The basic idea is that velocity of NIWs can be decomposed into

$$u + iv = e^{-if_0 t} LA(\varepsilon^2 t, x, y, z) \tag{2.1}$$

Here ε^2 represents the departure of the wave frequency from the inertial frequency f_0 and is required to be much smaller than unity to ensure that the internal waves are near-inertial. *A* represents the slow variation of NIW amplitude and phase on the time scale of $1/(\varepsilon^2 f_0)$ with L a differential operator defined by:

$$L = \frac{\partial}{\partial z}\left(\frac{f_0^2}{N^2}\frac{\partial}{\partial z}\right) \tag{2.2}$$

where $N(z)$ is the background buoyancy frequency.

Evolution of *A* is determined by the YBJ equation:

$$LA_t + J(\phi, LA) + i\frac{f_0}{2}\nabla^2 A + i\cdot(\beta\cdot y + \frac{1}{2}\varsigma)LA = 0 \tag{2.3}$$

(Young and Ben Jelloul 1997) where ϕ is stream function of background geostrophic flow, ∇^2 the horizontal Laplacian, β the planetary vorticity gradient, and $\varsigma = \nabla^2\phi$ the vorticity of background geostrophic flow.

In contrast to the ray tracing theory, the YBJ equation does not make spatial scale assumptions for background geostrophic flow and NIWs. This makes it suitable for examining the dispersion of wind-generated NIWs in presence of background geostrophic flow as the horizontal scale of midlatitude storms is typically an order of magnitude larger than that of mesoscale eddies. A formal demonstration for the validity of the YBJ equation for the studied problem in this chapter is provided in Appendix A.

2.2.2 *Vertical mode partition*

The vertical mode partition will serve an important analytical tool throughout the chapter. So it is briefly introduced here. Vertical modes are solutions of the Sturm-Liouville problem:

$$\frac{d}{dz}(\frac{1}{N^2}\frac{dF_n}{dz}) + \frac{1}{c_n^2}F_n = 0 \tag{2.4}$$

$$\frac{d}{dz}F_n = 0 \qquad z = 0, -H \tag{2.5}$$

where F_n is the *n*-th vertical mode, and c_n is the associated eigenvalue. Here the rigid-lid boundary condition is used for simplicity, which eliminates the barotropic mode. It should be noted that the wave frequency of barotropic waves is far from near-inertial, in

which case the YBJ equation is not applicable. Gill (1984) has shown that the barotropic response to wind forcing is almost instantaneous with weak currents. Therefore, the rigid-lid approximation is justified.

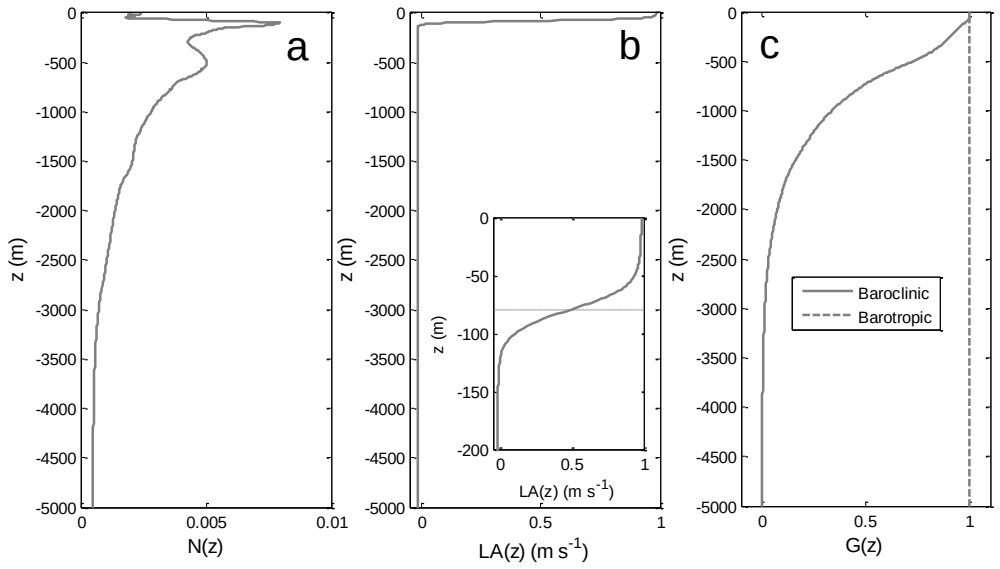

Figure 2.2 (a) Winter background buoyancy frequency in the midlatitude North Pacific (37.5°N, 159.5°W). (b) The initial vertical profile for LA. (c) The vertical profile of the baroclinic (solid) and barotropic (dashed) geostrophic flow.

The vertical modes in (2.4) and (2.5) form a complete and orthogonal basis. The completeness ensures that any function satisfying the boundary condition (2.5) can be represented completely by these vertical modes. And the orthogonality means:

$$\int_{-H}^{0} F_m F_n dz = \delta_{m,n} \tag{2.6}$$

where $\delta_{m,n}$ is the Kronecker delta.

The vertical modes are determined by the background stratification. Here the background stratification is derived from the temperature and salinity vertical profile in

the midlatitude North Pacific using the Word Ocean Atlas (Locarnini et al. 2010). Winter values are used as this is the season when energetic midlatitude storms are typically active (Alford 2003a) and thus give strong near-inertial energy input to the ocean. Figure 2.2a shows that the water column during this season is weakly stratified in the upper 60 m with almost a uniform potential density. The corresponding buoyancy frequency increases rapidly with increasing depth and peaks around 100 m with a maximum of 0.008 rad s^{-1}. There is a second peak centered around 500 m with a local maximum of 0.005 rad s^{-1}.

	E1	E2	E3
Φ (m^2 s^{-1})	4000	8000	16000
$\varsigma_{g,\min}$ (s^{-1})	-2.5×10^{-6}	-5.0×10^{-6}	-10.0×10^{-6}
$\hat{\beta}$	1.02	0.51	0.26

Table 2.1 Parameters used in the experiment E1, E2, and E3

2.2.3 *Initial condition for NIWs*

As horizontal scale of midlatitude storms is typically much larger than that of mesoscale eddies, NIWs are assumed to be horizontally homogeneous at initial. The initial vertical profile of LA is displayed in Figure 2.2b. Most of the near-inertial energy is confined to the mixed layer, rationalized by the *in situ* observations (Alford et al. 2012) and numerical simulations (Greatbatch 1984). To avoid projection onto the barotropic

mode, the vertical integration of LA is set to be zero, leading to a weak near-inertial current of ~0.02 m s^{-1} below 100 m.

2.2.4 *Idealized baroclinic geostrophic flow*

To facilitate theoretical analyses, an idealized baroclinic geostrophic flow is used here. The geostrophic flow is assumed to be zonally-directed and sinusoidal with the stream function:

$$\phi = -\Phi \cos(2\alpha y) G(z) \tag{2.7}$$

where Φ is a positive constant representing the intensity of geostrophic flow and $G(z)$ characterizes the vertical structure of geostrophic flow. The geostrophic flow would be barotropic and identical to that used by Balmforth et al. (1998) if $G(z)$ were set to unity. Here $G(z)$ is constructed by superposition of the barotropic and first baroclinic mode so that it is unity at the sea surface and decreases to zero at the bottom (Figure 2.2c), a more reasonable approximation for geostrophic flow in reality than a barotropic flow. Finally, β is taken as zero to isolate effects of the geostrophic flow on dispersion of NIWs following Balmforth et al. (1998). As NIWs are horizontally homogeneous at initial, the YBJ equation can be simplified as:

$$LA_t + i\frac{f_0}{2}A_{yy} + \frac{i}{2}\varsigma \cdot LA = 0 \tag{2.8}$$

where the advection term has been dropped as the geostrophic flow is zonally-directed and has no variation along the x-axis (The effect of advection term is discussed in Section 2.5.2). f_0 is taken as 10^{-4} rad s^{-1} in the following analyses. As ϕ is sinusoidal and A is horizontally homogeneous at initial, a periodic boundary condition is used at

the lateral boundaries $y = 0$ and $y = \pi / \alpha$ following Balmforth et al. (1998). Finally, a rigid-lid boundary condition is applied to filter out near-inertial barotropic current.

Three experiments are performed to analyze the dispersion of NIWs in the baroclinic geostrophic flow (See Table 2.1 for summary). In these three experiments, the intensity of geostrophic flow (Φ) varies from 4000 m^2s^{-1} to 16000 m^2s^{-1} but its horizontal structure is fixed with α^{-1}=80 km. The geostrophic vorticity minimum in the first experiment (referred to as E1 henceforth) is -2.5×10^{-6} s^{-1}, corresponding to a weak geostrophic flow. The geostrophic vorticity minimum in the experiment E2 and E3 increases to -5.0×10^{-6} s^{-1} and -10.0×10^{-6} s^{-1}, corresponding to a moderate and a strong geostrophic flow, respectively. Note that all these values are relevant to mesoscale eddies in the reality.

2.3 Dispersion of NIWs in a Baroclinic Geostrophic Flow

2.3.1 *Vertical dispersion*

An index D_V is introduced to measure the vertical dispersion of NIWs in the geostrophic flow. Here D_V is defined as $D_V = <Q^2> / <Q_0^2>$ where Q^2 (Q_0^2) is the (initial) energy of NIWs and < > denotes the horizontal average. For all three experiments (Figure 2.3), D_V decreases with time in the mixed layer and increases in the deep ocean (600-1000 m), consistent with a downward radiation of near-inertial energy from the mixed layer into the deep ocean.

The decay time of the near-inertial energy in the mixed layer varies considerably among the three experiments. In the experiment E1, less than 10% of near-inertial energy has radiated out of the mixed layer by 10 inertial periods (IPs) (Figure 2.3a). But this value increases to 30% and 70% in the experiment E2 and E3 (Figure 2.3b and 2.3c), suggesting that the effects of geostrophic flow are more significant with the stronger geostrophic vorticity.

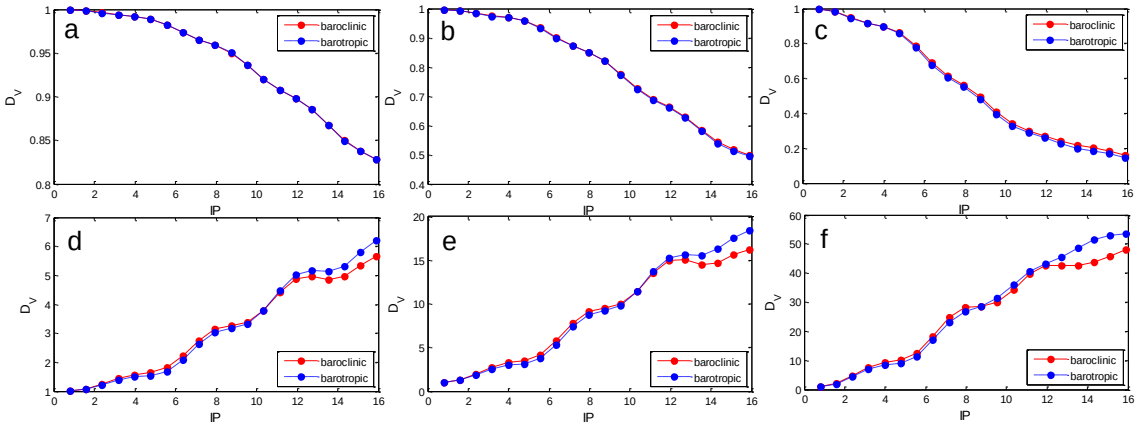

Figure 2.3 Vertical dispersion of NIWs. (a-c) display the vertical dispersion index D_V in the surface mixed layer for the experiment E1, E2, and E3 (from left to right), respectively. (d-f) are similar to (a-c) but for D_V in the deep ocean (600-1000 m). Note that the horizontal axis represents time in unit of inertial period and vertical axis ranges are different in different panels.

2.3.2 *Horizontal dispersion*

While the near-inertial energy is horizontally homogeneous at initial, evident inhomogeneity arises after several IPs. Throughout the water column, near-inertial energy tends to be concentrated in the negative-ς region (Figure 2.4). To further quantify the horizontal dispersion, we introduce an index D_H defined as

$D_H = std\{Q^2\} / < Q^2 >$ where $std\{\}$ denotes the standard deviation in the horizontal direction.

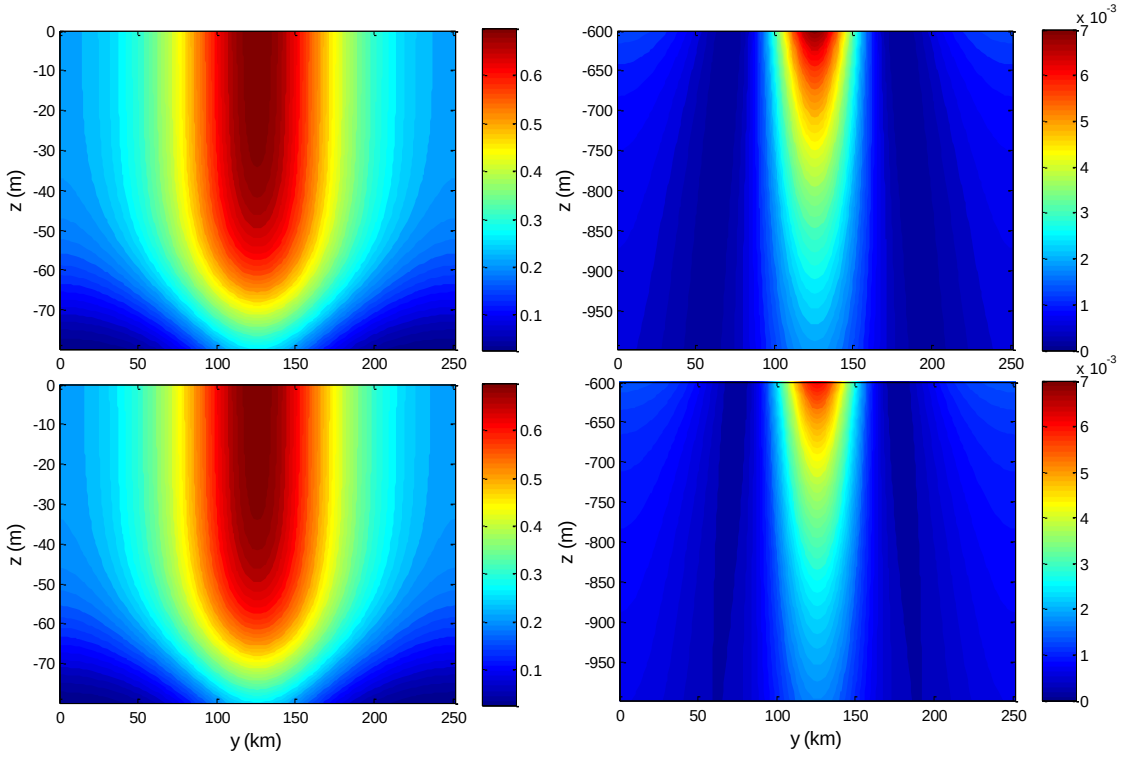

Figure 2.4 Snapshots of near-inertial kinetic energy (m^2s^{-2}) in the mixed layer (left panels) and thermocline (right panels) on 8 IPs for E2 (upper panels) and C2 (lower panels).

Compared with D_V, evolution of D_H is more complicated. Nevertheless, there is a common increase of D_H at the early stage (Figure 2.5). On one hand, the initial increase of D_H in the mixed layer corresponds to the concentration of near-inertial energy in the negative-ς region (Figure 2.4). On the other hand, initially increased D_H in the deep ocean results from the enhanced near-inertial energy around both the minimum and

maximum of ς (Figure 2.4). Such patterns are consistent with the findings reported by Balmforth et al. (1998). Variation of D_H becomes more rapid as the geostrophic flow intensity is elevated. For instance, it takes about 16 IPs for D_H in the surface mixed layer to increases from 0 to 0.5 in the experiment E1. But the time is reduced to about 9 IPs and 6 IPs in the experiment E2 and E3, respectively.

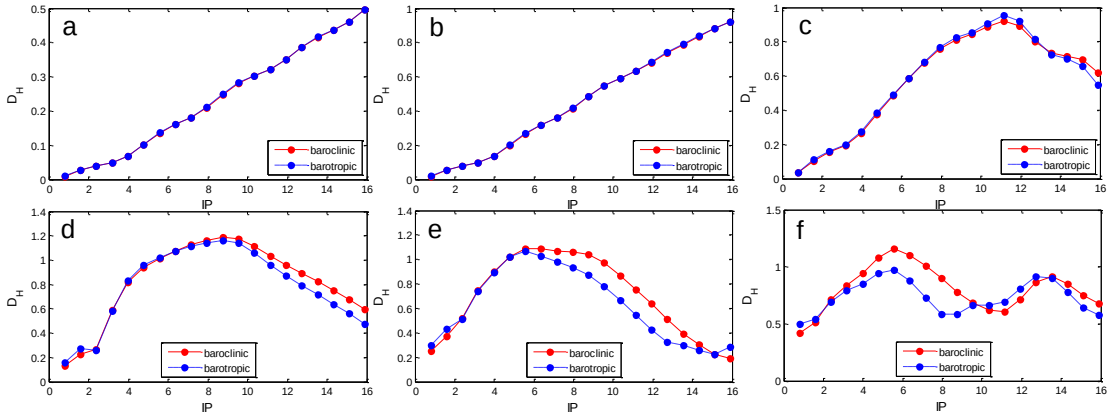

Figure 2.5 Similar to Figure 2.3 but for the horizontal dispersion of NIWs

2.3.3 *Comparisons with barotropic case*

To compare the dispersion of NIWs between barotropic and baroclinic flows, three control experiments (C1, C2, and C3) are performed. All the model configurations in the experiment C1, C2, and C3 are identical to those in E1, E2, and E3 except that $G(z)$ is set to unity i.e., a barotropic geostrophic flow. The evolution of D_V and D_H in the barotropic flow is generally comparable to that in the baroclinic flow (Figure 2.3 and 2.5), suggesting that baroclinic effect does not play an important role in modulating the dispersion of NIWs. It should be noted that the vorticity of the barotropic and baroclinic

flows is almost the same in the surface mixed layer (Figure 2.2c), which might account for the similarity of D_V and D_H in the mixed layer between the barotropic and baroclinic flows. However, the vorticity in the deep ocean is much weaker in the baroclinic flow than in the barotropic flow (Figure 2.2c) while the evolutions of D_V and D_H between these two cases are still similar there. Reasons for this similarity in the physical space will be examined in the following section.

2.4 Mechanisms for Dispersion of NIWs in a Baroclinic Geostrophic Flow

2.4.1 *Governing equations*

When the background geostrophic flow is barotropic, there are no interferences among different NIW modes so that each mode evolves independently (Balmforth et al. 1998). In this case, phase separation due to eigen-frequency difference among modes is the only mechanism responsible for the dispersion of near-inertial energy. In this section, we take the experiment E2 as an example to identify the corresponding mechanisms in the presence of a baroclinic geostrophic flow.

Expanding *A* by vertical modes yields:

$$\sum_{i=1}^{\infty}[\frac{\partial A_i}{\partial t} - i\frac{c_i^2}{2f_0}\frac{\partial^2 A_i}{\partial y^2} + 2i\alpha^2\Phi\cos(2\alpha y)G(z)A_i]F_i(z) = 0 \qquad (2.9)$$

where $F_i(z)$ is the *i*-th vertical mode and $LA = \sum_{i=1}^{\infty} A_i(y,t)F_i(z)$. Multiplying (2.9) by $F_m(z)$ and vertical integration yields:

$$(\frac{\partial A_m}{\partial t} - i\frac{c_m^2}{2f_0}\frac{\partial^2 A_m}{\partial y^2}) + 2i\alpha^2\Phi\cos(2\alpha y)\sum_{i=1}^{\infty} A_i \cdot r_{m,i} = 0 \qquad (2.10)$$

where $r_{m,i} = \int_{-H}^{0} G(z)F_m(z)F_i(z)dz$.

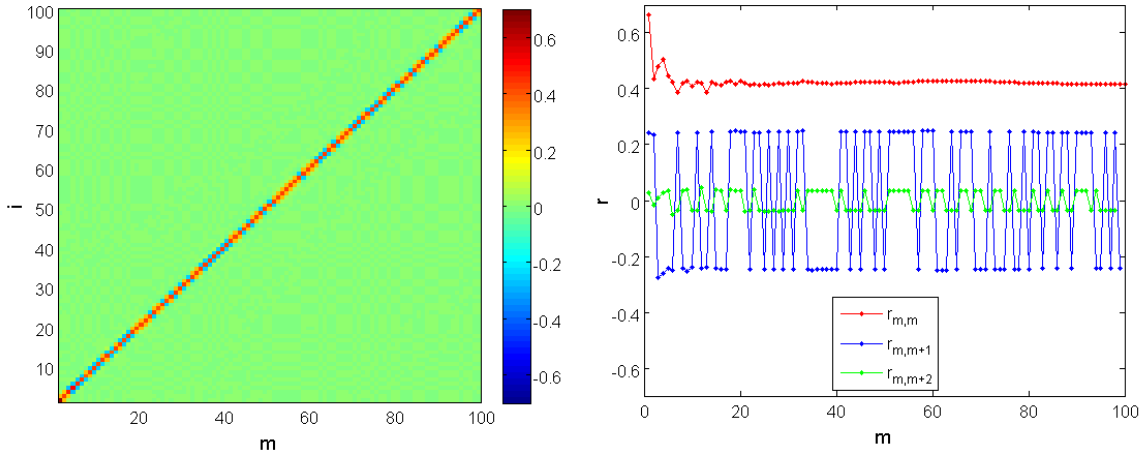

Figure 2.6 The dependence of $r_{m,i}$ on the vertical mode number.

Note that $r_{m,i}$ reduces to $\delta_{m,i}$ when the geostrophic flow is barotropic. For baroclinic background geostrophic flows, there are two important differences. First, the non-diagonal elements are not zero, corresponding to the interferences among NIWs of different vertical modes. Second, the diagonal elements are less than unity and exhibit a general decreasing trend with the increasing vertical mode number (Figure 2.6). For m>6, $r_{m,m}$ is almost unchanged and is approximately equal to 0.42. The smaller diagonal elements in the baroclinic flow are mainly due to the reduced vertical mean value of $G(z)$ compared to that in the barotropic flow (Figure 2.2c).

$A_m(y,t)$ can be further expanded by the Mathieu functions of even integer order (Balmforth et al. 1998):

$$A_m(y,t) = \sum_{j=0}^{\infty} B_{m,j}(t) ce_{2j}(\alpha y, q_m) \tag{2.11}$$

where $q_m = 2r_{m,m} f_0 \Phi / c_m^2$ and $ce_{2j}(\alpha y, q_m)$ are Mathieu functions of even integer orders.

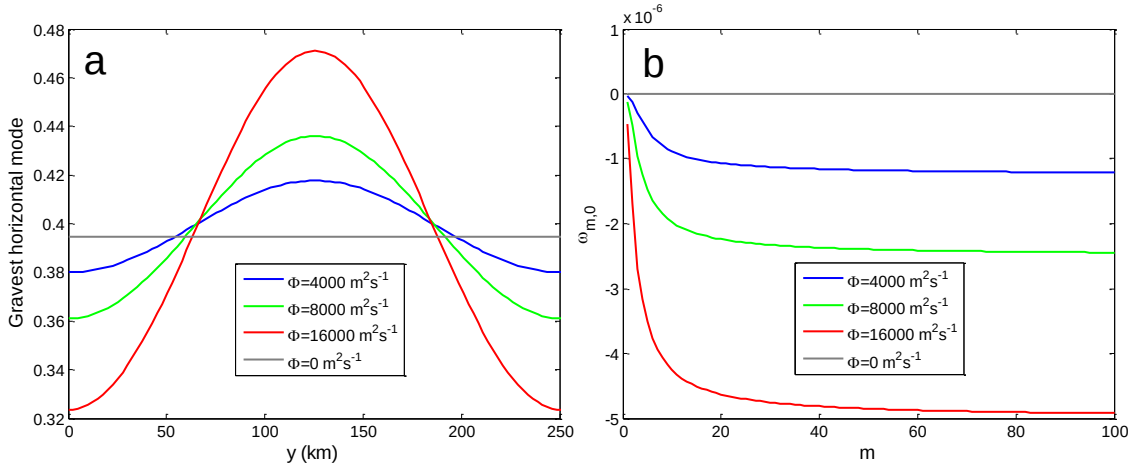

Figure 2.7 (a) The shape of the gravest horizontal mode of NIWs in the barotropic flows. (b) The dependence of eigen-frequencies $\omega_{m,0}$ on the vertical mode number for the gravest horizontal mode of NIWs in the barotropic flows.

Substituting (2.11) into (2.10) and some manipulations yield:

$$\sum_{j=0}^{\infty}[\frac{dB_{m,j}}{dt} + i\frac{\alpha^2 c_m^2}{2f_0} a_{m,j} B_{m,j}]ce_{2j}(\alpha y, q_m) + 2i\alpha^2 \Phi \cos(2\alpha y) \sum_{i \neq m} r_{m,i} \sum_{j=0}^{\infty} B_{i,j} ce_{2j}(\alpha y, q_i) = 0 \tag{2.12}$$

where $a_{m,j}$ is the eigenvalue associated with $ce_{2j}(\alpha y, q_m)$. Multiplying (2.12) by $ce_{2n}(\alpha y, q_m)$ and then taking a horizontal integration give:

$$\frac{dB_{m,n}}{dt} + i \cdot \omega_{m,n} B_{m,n} + \sum_{i \neq m} \sum_{j=0}^{\infty} s_{m,n,i,j} B_{i,j} = 0 \qquad (2.13)$$

where $s_{m,n,i,j} = 2i\alpha^2 \Phi \cdot r_{m,i} \int_0^{2\pi} \cos(2\alpha y) ce_{2j}(\alpha y, q_i) \cdot ce_{2n}(\alpha y, q_m) d(\alpha y)$ is the mode-mode interference coefficient and $\omega_{m,n} = \alpha^2 c_m^2 a_{m,n} / (2 f_0)$ corresponds to the eigen-frequency of each mode.

(2.13) is the controlling equation for the evolution of each NIW mode in the baroclinic geostrophic flow. Balmforth et al. (1998) provided a nice discussion on how the barotropic geostrophic vorticity influences the dispersion of NIWs. Here we briefly review their discussion as it helps understand the dispersion of NIWs in the presence of a baroclinic flow. In the barotropic case, effects of the geostrophic vorticity on NIW dispersion are measured by q_m. When q_m is zero, the Mathieu functions reduce to the sinusoidal functions. In this case, the initial near-inertial energy is totally projected onto the gravest horizontal mode $\cos(0 \cdot y)$. For the gravest horizontal mode, the eigen-frequencies $\omega_{m,0}$ for all the vertical modes are exactly zero. In this case, NIWs reduce to pure inertial oscillations with no horizontal and vertical energy redistribution. In presence of the geostrophic vorticity, q_m is nonzero. The geostrophic vorticity impresses its horizontal structure on the horizontal modes of NIWs. The gravest horizontal mode is no longer spatially constant (Figure 2.7a) so that the initial near-inertial energy is also projected onto the higher horizontal modes. As eigen-frequencies of horizontal modes differ from each other, phase separation leads to horizontal redistribution of near-inertial energy. It should be noted that the horizontal modes tend to be localized in the vicinity

of the vorticity minimum, corresponding to the concentration of near-inertial energy in the negative-ς region and analogous to the trapping of particles in the potential well in the quantum mechanics. Particularly, the trapping is more evident with the increased q_m (Figure 2.7a), leading to enhanced horizontal dispersion in the stronger geostrophic flow.

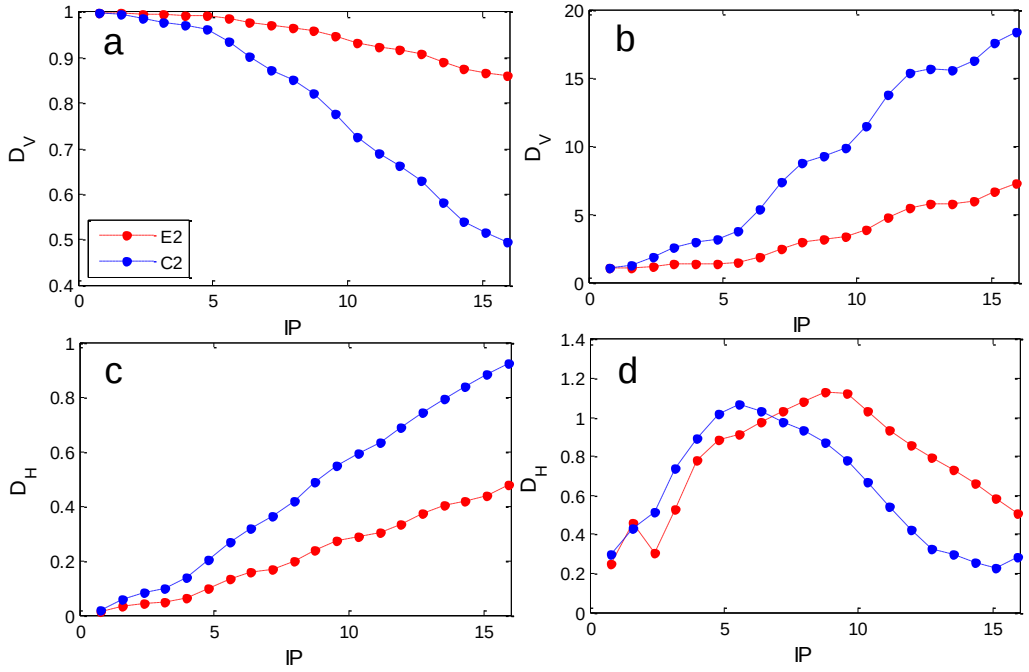

Figure 2.8 Dispersion of NIWs computed from the NMMI model in the experiment E2 (red) and C2 (blue). (a) D_V in the surface mixed layer, (b) D_V in the deep ocean, (c) D_H in the surface mixed layer, and (d) D_H in the deep ocean.

A non-zero q_m also modifies the eigen-frequency difference among vertical modes. Particularly, $\omega_{m,0}$ is no longer zero and differs among vertical modes (Figure 2.7b), resulting in a downward radiation of near-inertial energy from the surface mixed layer to

deep ocean due to the phase separation. As q_m increases, the eigen-frequency difference among vertical modes becomes larger, enhancing the vertical dispersion (Figure 2.7b).

Compared with the barotropic case, there are two major differences. First, mode-mode interferences of NIWs arise in the baroclinic geostrophic flow as the off-diagonal elements of $r_{m,i}$ are nonzero. Second, the shape and eigen-frequency of modes in the baroclinic case differ from those in the barotropic case due to the difference in $r_{m,m}$ (Note that q_m is proportional to $r_{m,m}$). In the following sections, we will examine how these differences may affect the dispersion of NIWs in baroclinic geostrophic flows.

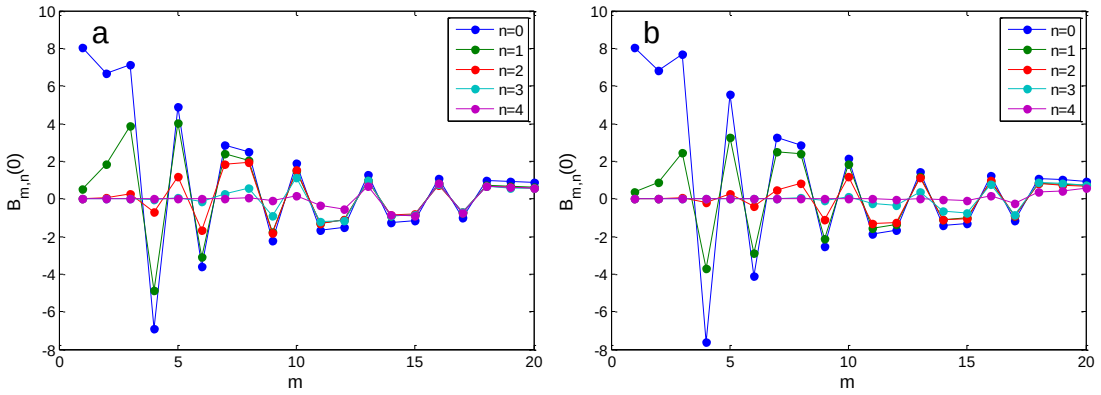

Figure 2.9 The initial modal projection coefficients $B_{m,n}(0)$ in the experiment (a) C2 and (b) E2. The dots in colors represent different horizontal modes.

2.4.2 *Eigen-frequency differences in a baroclinic geostrophic flow*

As mentioned above, both the eigen-frequency differences and mode-mode interferences may contribute to the dispersion of NIWs in baroclinic flows. In this section, we first examine the role of eigen-frequency differences in baroclinic flows and

compare it to that in barotropic flows. To isolate the contribution of eigen-frequency differences, the non-diagonal elements of $r_{m,i}$ are set to zero so that (2.13) can be simplified as:

$$\frac{dB_{m,n}}{dt} + i \cdot \omega_{m,n} B_{m,n} = 0 \tag{2.14}$$

(2.14) neglects the interferences among different modes of NIWs. For this sake, we refer to it as the non-mode-mode interference (NMMI) model. Figure 2.8 compares the evolution of D_V and D_H computed from the NMMI model in barotropic and baroclinic flows (Note that the NMMI model is equivalent to the YBJ equation in barotropic flows). For the NMMI model, the dispersion of NIWs is significantly weaker in baroclinic flows than in barotropic flows. For instance, the mixed-layer near-inertial energy in the experiment C2 decreases to 70% of its initial value after 10 IPs while more than 90% of the near-inertial energy still resides in the surface mixed layer in the experiment E2.

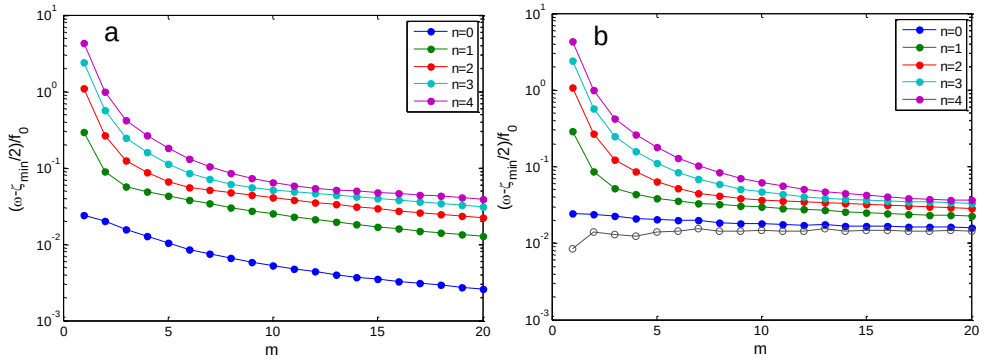

Figure 2.10 The normalized eigen-frequencies $(\omega_{m,n} - \varsigma_{\min}/2)/f_0$ in the experiment (a) C2 and (b) E2. The grey dots in (b) represent $(\varsigma_{\min}^{eff} - \varsigma_{\min})/(2f_0)$.

Both the shape and eigen-frequency of NIW modes differ between barotropic and baroclinic flows. The former controls the projection of near-inertial energy onto individual modes while the latter determines the dephasing time scale among modes. Both can influence the dispersion of NIWs. The initial modal projection coefficients exhibit minor differences between the experiment C2 and E2 due to their slightly differentiated horizontal modal shape (Figure 2.9). But their general patterns are very similar. In both cases, most of the near-inertial energy is projected onto the first two horizontal modes and the first ten vertical modes. However, unlike the modal projections, there are pronounced differences in the eigen-frequencies between barotropic and baroclinic flows especially for the gravest horizontal mode (Figure 2.10). The eigen-frequencies for the gravest horizontal mode become significantly higher in the baroclinic case than in the barotropic case, reducing the difference between $\omega_{m,0}$ and $\omega_{m,1}$ and thus weakening the horizontal dispersion. Furthermore, the eigen-frequencies for the gravest horizontal mode become less sensitive to the changes of vertical mode number in the baroclinic case than in the barotropic case, which slows down the vertical dispersion of NIWs.

The reasons for the reduced eigen-frequency differences in the baroclinic flow can be understood by introducing an effective geostrophic flow intensity $\Phi_m^{eff} = r_{m,m}\Phi$. According to (2.12) and (2.13), the eigen-frequencies are determined by $q_m = 2r_{m,m} f_0 \Phi / c_m^2 = 2 f_0 \Phi_m^{eff} / c_m^2$. Therefore, it is Φ_m^{eff} rather than Φ that eventually determines the impact of background geostrophic flows on eigen-frequencies of NIW

modes. As $r_{m,m}$ is smaller than unity in baroclinic flows due to the weaker ς below the surface mixed layer (Figure 2.6b), Φ_m^{eff} in baroclinic flows is always smaller than Φ. Correspondingly, the eigen-frequency differences in a baroclinic flow are always smaller than those in a barotropic flow with the same surface intensity. This explains why the dispersion of NIWs derived from the NMMI model is significantly weaker in the experiment E2 than in the experiment C2.

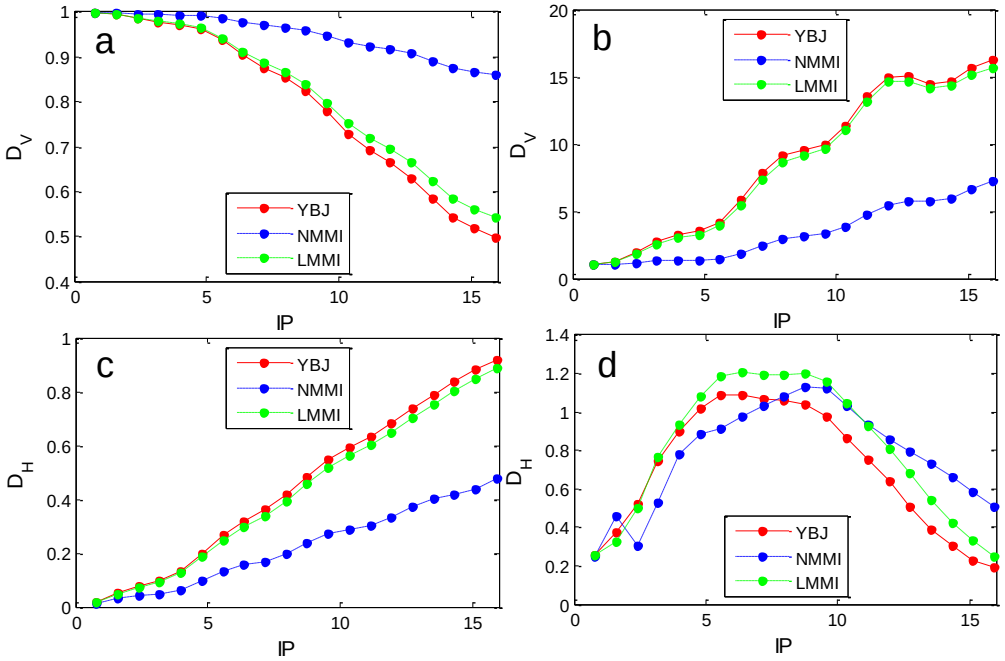

Figure 2.11 Dispersion of NIWs in the experiment E2 computed from the YBJ equation (red), the NMMI model (blue), and the LMMI model (green). (a) D_V in the surface mixed layer, (b) D_V in the deep ocean, (c) D_H in the surface mixed layer, and (d) D_H in the deep ocean.

Finally, it should be noted that the dispersion of NIWs in the experiment E2 is significantly underestimated by the NMMI model. In the experiment E2, evolution of

D_V and D_H computed from the NMMI model is much slower than that from the YBJ equation (Figure 2.11), suggesting that the mode-mode interferences may play an important role in the dispersion of NIWs in baroclinic flows. This will be examined in the next section.

2.4.3 *Mode-mode interferences of NIWs in a baroclinic geostrophic flow*

As demonstrated in Section 2.4.2, the mode-mode interferences make a significant contribution to the dispersion of NIWs in baroclinic flows. In the modal space, the interferences can be either locally or non-locally. In this section, we will examine which kind of interferences plays a key role in the dispersion of near-inertial energy.

The mode-mode interferences result from the non-zero values of $r_{m,i}$ for $m \neq i$. As *G* is constructed using the barotropic mode and the first baroclinic mode, the value of $r_{m,i}$ is expected to reduce rapidly with the increased value of $|m-i|$. In fact, $|r_{m,i}|$ is about 0.2 for $|m-i|=1$ but decreased to a value smaller than 0.04 for $|m-i|\geq 2$ (Figure 2.6), suggesting that the nonlocal interferences in the vertical-modal space may make negligible contribution to the dispersion of NIWs. In this case, (2.13) can be approximated as:

$$\frac{dB_{m,n}}{dt} + i \cdot \omega_{m,n} B_{m,n} + \sum_{j=0}^{\infty} s_{m,n,m-1,j} B_{m-1,j} + \sum_{j=0}^{\infty} s_{m,n,m+1,j} B_{m+1,j} = 0 \qquad (2.15)$$

(2.15) excludes the nonlocal mode-mode interferences in the vertical modal space while both the local and nonlocal interferences in the horizontal modal space are retained. The relative importance between the local and nonlocal interferences in the horizontal

modal space is determined by the matrixes $s_{m,n,m-1,j}$ and $s_{m,n,m+1,j}$ for a particular m. Figure 2.12 displays $s_{m,n,m+1,j}$ for various values of m (The pattern of $s_{m,n,m-1,j}$ is very similar and thus not shown). For any m, $s_{m,n,m-1,j}$ and $s_{m,n,m+1,j}$ are approximately tridiagonal matrixes, suggesting that the local interferences in the horizontal modal space may dominate. Therefore, (2.15) can be further simplified as:

$$\frac{dB_{m,n}}{dt} + i \cdot \omega_{m,n} B_{m,n} + \sum_{j=n-1}^{n+1} s_{m,n,m-1,j} B_{m-1,j} + \sum_{j=n-1}^{n+1} s_{m,n,m+1,j} B_{m+1,j} = 0 \qquad (2.16)$$

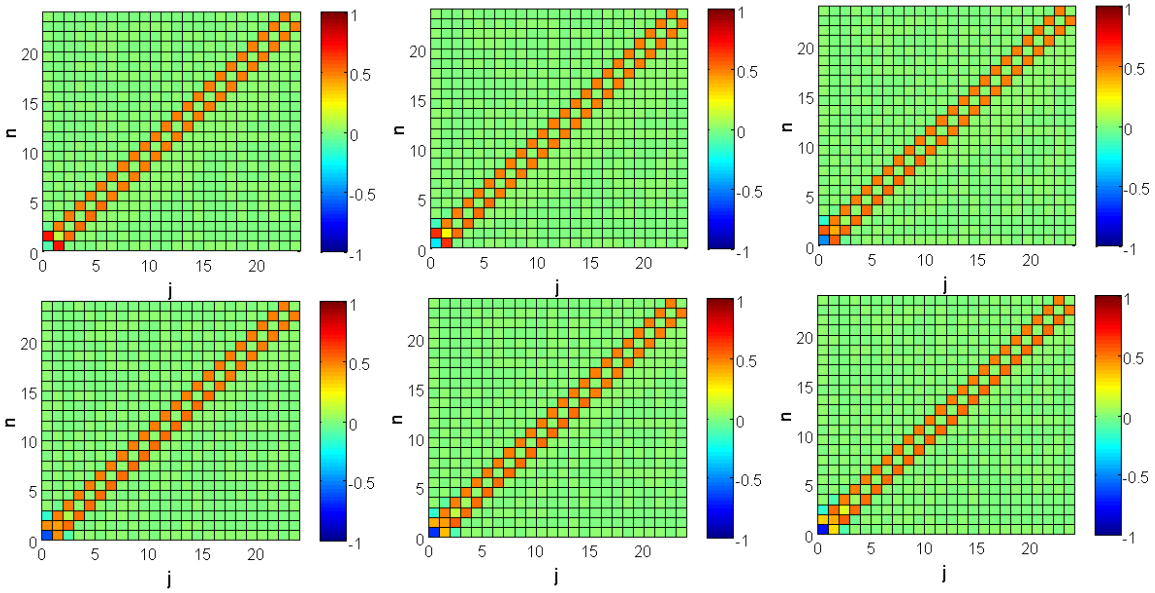

Figure 2.12 The horizontal mode-mode interference coefficient matrix $s_{m,n,m+1,j}$ for m=1, 2, 3, 4, 5, and 6.

Compared to (2.13), (2.16) only includes interferences locally in both the horizontal and vertical modal space and is thus referred to as the local mode-mode-interference (LMMI) model. It is simplified version of the YBJ equation and can be numerically

solved using a Runge-Kutta method. The validity of LMMI model is evaluated based on the comparisons with the solution derived from the YBJ equation. As can be seen from Figure 2.11, the evolution of D_V and D_H computed from the LLMI model agrees well with that from the YBJ equation. Particularly, the difference is less than 20% in the first 10 IPs. Therefore, the local mode-mode interferences play a dominant role in the dispersion of near-inertial energy compared to the nonlocal mode-mode interferences.

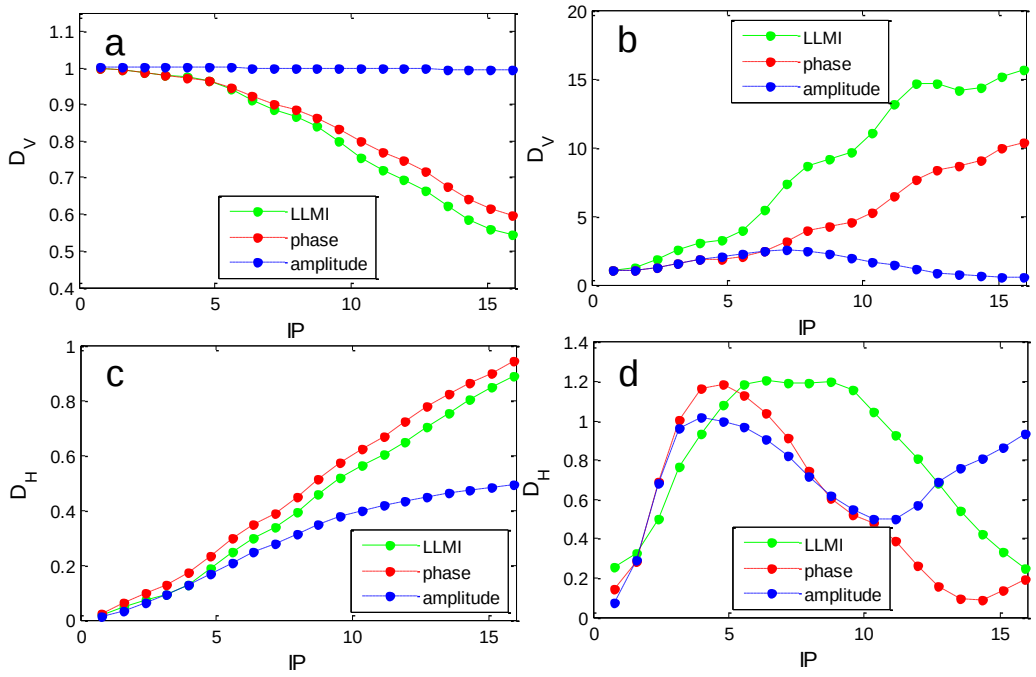

Figure 2.13 The dispersion of NIWs in the experiment E2 computed from the LLMI model (blue). The red and blue dashed lines represent the contribution from the phase separation and energy exchange, respectively. (a) D_V in the surface mixed layer, (b) D_V in the deep ocean, (c) D_H in the surface mixed layer, and (d) D_H in the deep ocean.

The local mode-mode interactions modify the evolution of phase and amplitude of individual modes. In this case, the phase separation may not be the only mechanism

responsible for the dispersion of NIWs as the energy exchange among modes is also able to redistribute the near-inertial energy in the physical space. For instance, trapping of near-inertial energy in the negative-ς region is more evident for the lower horizontal modes. When energy was transferred from the low horizontal modes to the high horizontal modes, the near-inertial energy would be distributed more uniformly in the horizontal direction and vice versa.

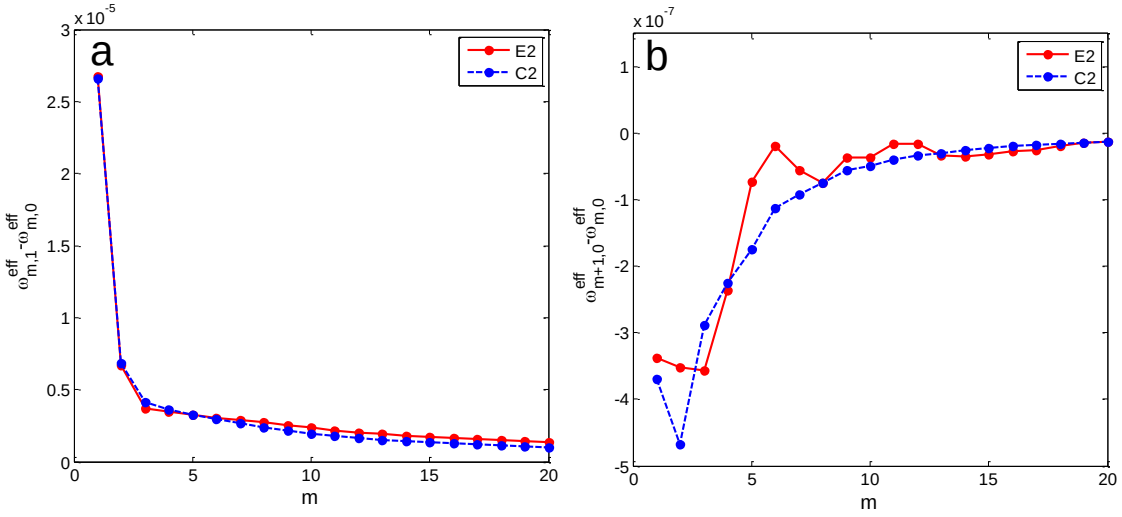

Figure 2.14 Effective eigenfrequency differences in the experiment E2 (red) and C2 (blue): (a) $\omega_{m,1}^{eff} - \omega_{m,0}^{eff}$ and (b) $\omega_{m+1,0}^{eff} - \omega_{m,0}^{eff}$.

To examine the relative importance between the phase separation and energy exchange among modes, we re-compute D_V and D_H using the LMMI model but fixing either the amplitude or phase of $B_{m,n}$ as its initial value. The energy exchange has little influence on the dispersion of NIWs in the surface mixed layer (Figure 2.13). Fixing the

amplitude does not make appreciable impacts on the evolution of D_V and D_H. In the deep ocean, the energy exchange among modes becomes somewhat more important (Figure 2.13). Nevertheless, it is still the phase separation that dominates the dispersion of NIWs.

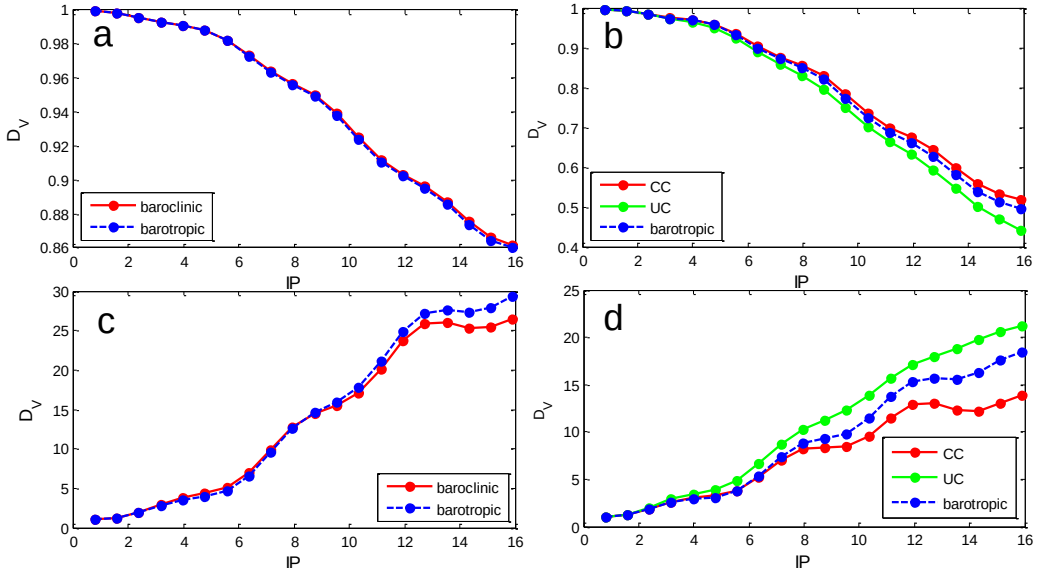

Figure 2.15 Vertical dispersion of NIWs in the surface mixed layer (a and b) and the deep ocean (c and d). The left panels display the values in the barotropic (blue) and baroclinic (red) flows using the autumn stratification. The right panels show the values in geostrophic flows of various vertical structures: counter current (red), undercurrent (green), and barotropic flow (blue).

Finally, we introduce an effective eigen-frequency defined as:

$$\omega_{m,n}^{eff} = \frac{\Delta\theta_{m,n}}{\Delta t} \tag{2.17}$$

where $\theta_{m,n} = \arg(B_{m,n})$ is the phase for each mode and Δt is the time interval. Here $\Delta\theta_{m,n}$ is computed as the difference between the beginning and the end of the

simulations based on the LMMI model. However, using different time intervals does not make any substantial impact on $\omega_{m,n}^{eff}$.

$\omega_{m,n}^{eff}$ measures the combined contribution to the phase evolution from the eigen-frequency and local mode-mode interferences. While the eigen-frequency difference is significantly smaller in the experiment E2 than C2, the effective eigen-frequency differences are comparable between the two cases (Figure 2.14). The reduced eigen-frequency difference in baroclinic geostrophic flows is largely compensated by the additional phase separation resulting from the local mode-mode interferences, leading to a dephasing time scale comparable to that in barotropic flows. As the phase separation among modes is the major mechanism responsible for the dispersion of NIWs in barotropic and baroclinic flows, this explains why the dispersions of NIWs in barotropic and baroclinic flows are similar to each other.

2.5 Discussion

2.5.1 *The generosity of the compensation effect*

In Section 2.4.3, we demonstrated that the reduced eigen-frequency difference in the baroclinic geostrophic flow is largely compensated by the phase separation resulting from the local mode-mode interferences. However, it remains uncertain whether the compensation effect is a universal feature in baroclinic geostrophic flows. We re-run the experiments by modifying the model configurations and examine the sensitivity of the compensation effect to the background stratification and vertical structure of geostrophic flow.

We first change the background stratification to an autumn profile with a much shallower mixed layer (~30 m). In this case, the eigenvalues and vertical modes are slightly modified while the initial projection of NIWs on the vertical modes is significantly changed as the initial near-inertial energy is now confined to a much shallower region (not shown). However, the vertical dispersion of NIWs in the baroclinic case is still very similar to that in the barotropic case (Figure 2.15a and c). Therefore, the compensation effect does not seem to be sensitive to the variation of background stratification.

We then change the vertical structure of the geostrophic flow. The baroclinic geostrophic flow is reconstructed by the barotropic, the first and the second baroclinic mode with $G(z)$ fixed as unity (zero) at the surface (bottom). We consider two cases that are not rare in the reality. In the first case, there is a counter current below the thermocline with $G(z) = -0.37$ at 1100 m. In the second case, $G(z)$ peaks at 740 m with a maximum of 2.4, corresponding to a strong undercurrent. In both the two cases, the vertical dispersion of NIWs exhibit minor differences between the barotropic and baroclinic geostrophic flows (Figure 2.15b and d), suggesting that the vertical structure of the geostrophic flow does not have a significant influence on the compensation effect either.

Results above suggest that the compensation effect appears to be a general feature for dispersion of NIWs in baroclinic geostrophic flows. This compensation effect in the modal space may be a constraint imposed by the local similarity in the physical space. In our simulations, the barotropic and baroclinic geostrophic flows are almost the same in

the mixed layer. Therefore, the dispersion of NIWs in the surface mixed layer is expected to be very similar between the barotropic and baroclinic case. Note that the dispersion of NIWs in the surface mixed layer is controlled by the phase separation (Figure 2.13). In this case, the local mode-mode interferences are required to compensate the reduced eigen-frequency difference in the baroclinic geostrophic flow, ensuring a dephasing time scale comparable to that in the barotropic case. As the phase separation also dominates the dispersion of NIWs in the deep ocean (Figure 2.13), this further leads to similar dispersion of NIWs there between the barotropic and baroclinic case.

2.5.2 *The β effect and Doppler shift*

To facilitate theoretical analyses, β-effect is removed from our idealized model. The relative importance between the geostrophic vorticity and β-effect can be measured as $\hat{\beta} = \beta/(2\alpha^3\Phi)$ (Balmforth and Young 1999). When $\hat{\beta}$ is much smaller than unity, β effect plays a negligible role. The associated values of $\hat{\beta}$ for the experiment E1, E2, and E3 are 1.02, 0.51, and 0.26, respectively (Table 2.1). Therefore, β-effect is unlikely to have a substantial impact on the simulation results in the experiment E2 and E3.

Another deficiency in our idealized geostrophic flow is the exclusion of the Doppler shift. As the geostrophic flow is assumed to be two dimensional, the advection term vanishes. Klein and Smith (2001) analyzed the dispersion of NIWs in a turbulent baroclinic mesoscale eddy field. Similar to the results here, they also found that the dispersions of NIWs are comparable between barotropic and baroclinic flows with the same surface intensity. In particular, keeping and dropping advection terms only leads to

slight disparities with negligible effects on the spatial distribution of near-inertial energy (Klein et al. 2004). Therefore, conclusions derived from the idealized geostrophic flow may also hold qualitatively in the reality.

2.6 Summary

Horizontal and vertical dispersion of NIWs in an idealized baroclinic geostrophic flow is analyzed based on the YBJ equation. The corresponding background geostrophic stream function is taken the form of $\phi = -\Phi\cos(2\alpha y)G(z)$, where $G(z)$ is constructed by superposition of the barotropic and first baroclinic mode so that it is unity at the sea surface and decreases to zero at the bottom. The major conclusions are summarized as follows:

(1) In the presence of baroclinic geostrophic flows, different modes of NIWs do not evolve independently. Local interferences in modal space play an important role in the dispersion of NIWs while the effects of nonlocal mode-mode interferences are negligible. A simplified version of the YBJ equation, the LLMI model, is constructed by dropping the nonlocal mode-mode interference terms and shows good skills in simulating the NIW dispersion in the baroclinic geostrophic flow.

(2) The dispersion of NIWs in the baroclinic geostrophic flow is achieved by both phase separation and energy exchange among modes with the former playing a more dominant role. The phase separation results from the eigen-frequency difference and local mode-mode interferences. Compared with the barotropic

case, the eigen-frequency difference is reduced in the baroclinic flow due to the weaker geostrophic flow below the surface mixed layer. But the reduction is largely compensated by the contribution from the local mode-mode interferences, leading to a dephasing time scale comparable to that in the barotropic case. Correspondingly, the dispersions of NIWs in barotropic and baroclinic flows are similar to each other.

CHAPTER III

ENERGY EXCHANGE BETWEEN NEAR-INERTIAL INTERNAL WAVES AND MESOSCALE EDDIES: ROLE OF WAVE CAPTURE

3.1 Introduction

Energy exchange between mesoscale eddies and internal waves is proposed to play an important role in the ocean energy budget (Ferrari and Wunsch 2009). On one hand, forward energy cascade from mesoscale eddies to small-scale internal waves can be a significant energy sink of mesoscale eddies (Polzin 2010). On the other hand, it may also provide a significant energy source for internal wave field. Several mechanisms have been proposed for energy exchange between mesoscale eddies and internal waves (e.g., Bretherton 1966; Müller 1976; Bühler and McIntyre 2005; Vanneste 2013). Müller (1976) analyzed energy exchange between geostrophic flows and internal waves based on the radiation-balance equation and suggested that relaxation effects of internal waves play a key role in permanent energy transfer from geostrophic flows to internal waves. He further evaluated the form of internal-wave induced diffusion by assuming that the relaxation effect dominates the propagation effect. This is termed the local limit approximation (Müller 1976) as the distortion of internal wave equilibrium spectrum by geostrophic flows is local in time, physical and wavenumber space. In such a scenario, the diffusion is linearly dependent on the relaxation time scale and internal wave energy. The predicted horizontal ν_h and vertical viscosity ν_v are ~7 m^2s^{-1} and ~0.4 m^2s^{-1}, respectively.

The relaxation time scale of internal waves is typically of O(1-10 days) (Frankignoul 1976; Olbers 1976; Brown and Owens 1981). It is not obvious that the relaxation effect should dominate the propagation effect. Actually, *in situ* observations (Frankignoul 1976; Ruddick and Joyce 1979; Frankignoul and Joyce 1979; Polzin 2010) do not always support the values predicted by Müller (1976), casting doubt on the validity of the local limit approximation. For instance, Ruddick and Joyce (1979) found in Polymode currentmeter data that the magnitude of ν_v is less than 0.02 m^2s^{-1} with some uncertainty in its sign in regions of low-eddy activity.

At the other limit of the radiation-balance equation where the relaxation effect is negligible compared to the propagation effect, evolution of internal waves in the geostrophic flows can be described by ray-tracing equations (Lighthill 1978):

$$\frac{dk}{dt} = -U_x k - V_x l \tag{3.1}$$

$$\frac{dl}{dt} = -U_y k - V_y l \tag{3.2}$$

$$\frac{dm}{dt} = -U_z k - V_z l - \Omega_z \tag{3.3}$$

where $\vec{K} = (k, l, m)$ is the wavenumber vector of internal waves, (U, V) background geostrophic currents, Ω the internal wave dispersion relation, and d/dt the time rate of change along ray paths. Despite the complexity of a three-dimensional geostrophic flow, Bühler and McIntyre (2005) proposed a simple characterization of interactions between the internal waves and geostrophic flow. Simply put, evolution of the wave number vector is determined by the Okubo-Weiss (OW) parameter (Provenzale 1999) defined as

strain variance minus vorticity variance, provided that background buoyancy frequency and gradient of geostrophic flows is approximately constant along the ray (See Appendix B for detail). When the OW parameter is negative, the horizontal wave vector rotates and its magnitude oscillates. The geostrophic flow has no permanent influences on internal waves in linear sense. However, azimuth of horizontal wave vector asymptotically points to a direction determined by the geostrophic velocity gradient alone when the OW parameter is positive. In that direction, the magnitude of wavenumber exhibits exponential growth at large time (See appendix B for details). As the group velocity of internal waves decreases with increasing wavenumber magnitude (Bühler and McIntyre 2005), internal waves will be eventually captured by the geostrophic flow, which is known as the wave capture (Bühler and McIntyre 2005) or shrinking catastrophe (Jones 1969). In the wave capture scenario, there is a negative stress-strain correlation for linear internal waves (i.e., a positive horizontal viscosity in a flux-gradient closure) as a consequence of wave action conservation, leading to a permanent energy transfer from the geostrophic flow to the internal wave field (See appendix B for details).

The local limit approximation and wave capture can be treated as two limiting cases in the framework of radiation-balance equation. Due to limited observations, it remains uncertain which one plays a more important role in energy exchange between mesoscale eddies and internal waves. The mooring array deployed in the northern Gulf of Mexico during the Gulf Integrated Spill Research Consortium (GISR) project (http://www.gulfbase.org/project/view.php?pid=gomisrcg) provides a unique

opportunity to study the energy exchange between mesoscale eddies and internal waves (Figure 3.1). In particular, the relative importance of these two mechanisms may be evaluated using the observed data. Furthermore, to examine whether the conclusions derived from the mooring measurements are applicable throughout the Gulf of Mexico, a high-resolution numerical simulation in the Gulf of Mexico is performed to validate the result using the Regional Ocean Modelling System (ROMS) (Shchepetkin and McWilliams 2005). This chapter is organized as follows. Data and methodology are given in section 3.2. Mooring data analyses and numerical simulation results are presented in section 3.3 and 3.4, respectively. Conclusions and discussion are provided in section 3.5.

3.2 Data and Methodology

3.2.1 *Mooring data*

In July of 2012, six moorings, henceforth referred to as M1, M2, and etc., were deployed in the northern Gulf of Mexico (Figure 3.1a) as a part of oil-spill studies conducted by GISR. Each mooring was equipped with a Rowe-Deines Instruments (RDI) Broadband 75 kHz Long-Ranger ADCP with 16 m bin spacing. Only M1, M2, and M3 are used here as they are close to each other and all sample the water column between 245-450 m (Figure 3.1b), making it possible to compute the horizontal velocity gradient of mesoscale eddies. There were also several currentmeters below the ADCPs to measure the velocity in the deep ocean (Figure 3.1b). The velocity measured by ADCPs and currentmeters is recorded every half hour. It is then interpolated onto a 5-m regularly

spaced vertical grid. 26 CTD profiles were collected during the retrieve of moorings. They are used to compute the stratification.

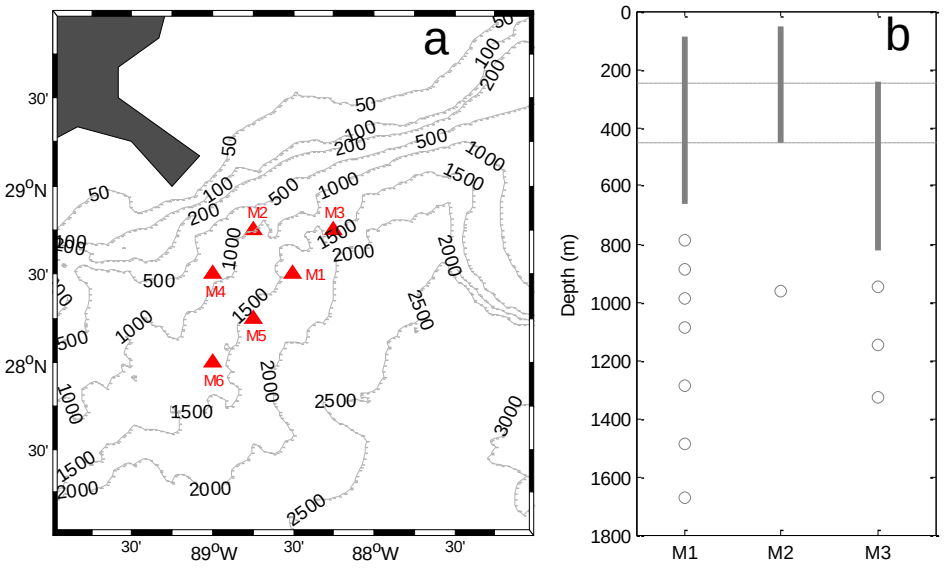

Figure 3.1 (a) The geography of the mooring array with contours denoting the topography; (b) The velocity measurements on M1, M2, M3. The solid lines represent the sampling interval of ADCP while the circles denote the currentmeters. The dash lines bound the region analyzed in this study.

The instruments were deployed for one year. Mooring layover is minimal with a root-mean-square (rms) vertical excursion of 0.36 m for M1, 0.30 m for M2, and 0.30 m for M3. In particular, the ADCPs always stayed within 1 m of their target depth except that M1 underwent a rise of 3 m in the end of June, 2013. But this is far too little to have a significant impact on the measurements.

3.2.2 *Isolating the internal waves and mesoscale eddies*

The horizontal velocity of internal waves is obtained by high-pass filtering the ADCP records. The half-power point is chosen to be 0.8 *f*, following Brown and Owens

(1981). Flow of mesoscale eddies is attained by low-pass filtering the data with a half-power point of 6 inertial periods (IPs). Changing the half-power point from 2 to 10 IPs makes no appreciable impacts. The spectra and co-spectra of internal-wave horizontal velocity are computed with 3-IP-long transform intervals. The 3-IP interval is used because it is long enough to resolve low-frequency internal waves (e.g., NIWs) but is short enough so that the geostrophic flows remain almost unchanged.

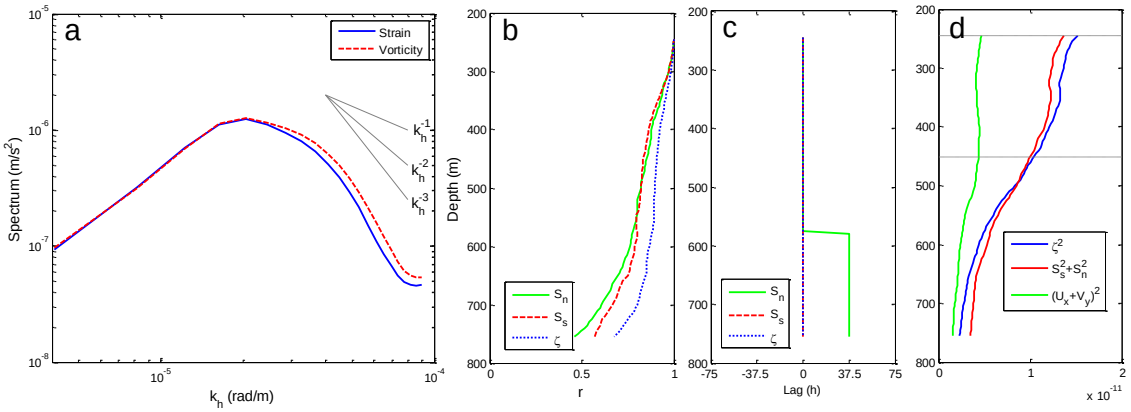

Figure 3.2 (a) The time-mean (July 2012-July 2013) wavenumber spectrum for vorticity and strain computed from sea level anomaly in the northern Gulf of Meixco (22°N-29°N, 95°W-87°W) obtained from the Ssalto/Duacs multimission altimeter products of AVISO. Here the strain spectrum is defined as the sum of the spectrum of $S_n = U_x - V_y$ and that of $S_s = U_y + V_x$. (b) The lag correlation of S_n, S_s, and ς at 245 m to those at various depths. (c) The time lag in hour associated with the highest correlation. (d) The time-mean ς^2, $S_n^2 + S_s^2$, and $(U_x + V_y)^2$.

The horizontal velocity gradient of mesoscale eddies is computed from the velocity difference among M1, M2, and M3. Due to the finite distance (~40 km) between moorings, velocity gradient of mesoscale eddies cannot be fully resolved. In the northern Gulf of Mexico (22°N-29°N, 95°W-87°W), the deformation radius L_D of first baroclinic

mode ranges from 30-50 km, suggesting that the strain and vorticity associated with mesoscale eddies would have a dominant wavelength of 180-300 km ($2\pi L_D$). This is confirmed by the strain and vorticity spectra computed from the 1/3° merged sea level anomaly obtained from the Ssalto/Duacs multimission altimeter products of AVISO, which exhibit a pronounced peak around wavelength of 300 km (Figure 3.2a). Therefore, a 40-km horizontal resolution should be able to resolve bulk of the mesoscale strain and vorticity variance.

The mesoscale eddies are vertically coherent in the upper 200-800 m (Figure 3.2b and c). But their strength decreases roughly linearly as depth increases (Figure 3.2d), implying the dominance of the first baroclinic mode. In the upper 245-450 m, the mesoscale eddies are associated with a root-mean-square (RMS) Rossby number of 0.05. The strength of convergence/divergence is much smaller than that of strain or vorticity (Figure 3.2d), consistent with the quasi-geostrophic regime (Pedlosky 1987). In the deeper region, the ratio of convergence/divergence to strain or vorticity is increased. This may be partly due to the errors resulting from vertical interpolation (Figure 3.1b) or ageostrophic motions induced by topography. In Section 3.3, the analysis will be confined to 245-450 m where the geostrophic approximation is appropriate

3.2.3 *Azimuth of horizontal wave vector of internal waves*

To assess the role of wave capture, estimates for the azimuth θ of horizontal wave vector are required. According to the polarization relation of internal waves, θ can be estimated from the relation:

$$\tan 2\theta = 2\,\mathrm{Re}\{P_{uv}\}/(P_{uu} - P_{vv}) \tag{3.4}$$

where P_{uv} is the cospectrum between zonal u and meridional velocity v of internal waves, P_{uu} (P_{vv}) the power spectrum for u (v), and $\mathrm{Re}\{\bullet\}$ the real operator.

Errors in ADCP records may contaminate the azimuth estimates. The noise level of ADCP records is typically of O(0.01 m s^{-1}). Based on Monte Carlo simulations in which the noise is assumed to be normally distributed with a standard deviation of 0.01 m s^{-1}, the resulted rms error in θ ranges from 12^o to 20^o and becomes larger with the increasing wave frequency (not shown). The larger RMS error for high-frequency internal waves is mainly due to their weak currents, leading to a low signal/noise ratio.

3.2.4 *Energy exchange between mesoscale eddies and internal waves*

Energy exchange between mesoscale eddies and internal waves is determined by the relation between internal-wave stress and velocity gradient of mesoscale eddies. In particular, previous studies indicated that internal wave-induced horizontal viscosity makes a major contribution to the energy exchange (Polzin 2010). As the vertical velocity estimates are not available, we will focus on the relation between internal-wave stress and horizontal velocity gradient of mesoscale eddies. The corresponding energy transfer rate is (Polzin 2010):

$$\varepsilon = -(<uu> - <vv>)\frac{S_n}{2} - <uv> S_s \qquad (3.5)$$

where $S_n = U_x - V_y$ and $S_s = U_y + V_x$ are the normal and shear components of mesoscale eddy strain, respectively, and $<\bullet>$ represents the running mean over 3-IP interval. In aid of the Parseval theorem, (3.5) can be re-expressed in the frequency domain:

$$\varepsilon = -\int (P_{uu} - P_{vv}) d\omega \frac{S_n}{2} - \mathrm{Re}\{\int P_{uv} d\omega\} S_s \qquad (3.6)$$

3.2.5 High-resolution numerical simulations

A high-resolution numerical simulation is configured over the entire Gulf of Mexico using ROMS. The horizontal resolution of ROMS is set at 1 km with 60 vertical layers. The nonlocal K-profile parameterization (Large et al. 1994) is used to parameterize vertical mixing. A harmonic horizontal mixing is used with both diffusivity and viscosity values set at 5 m^2s^{-1}.

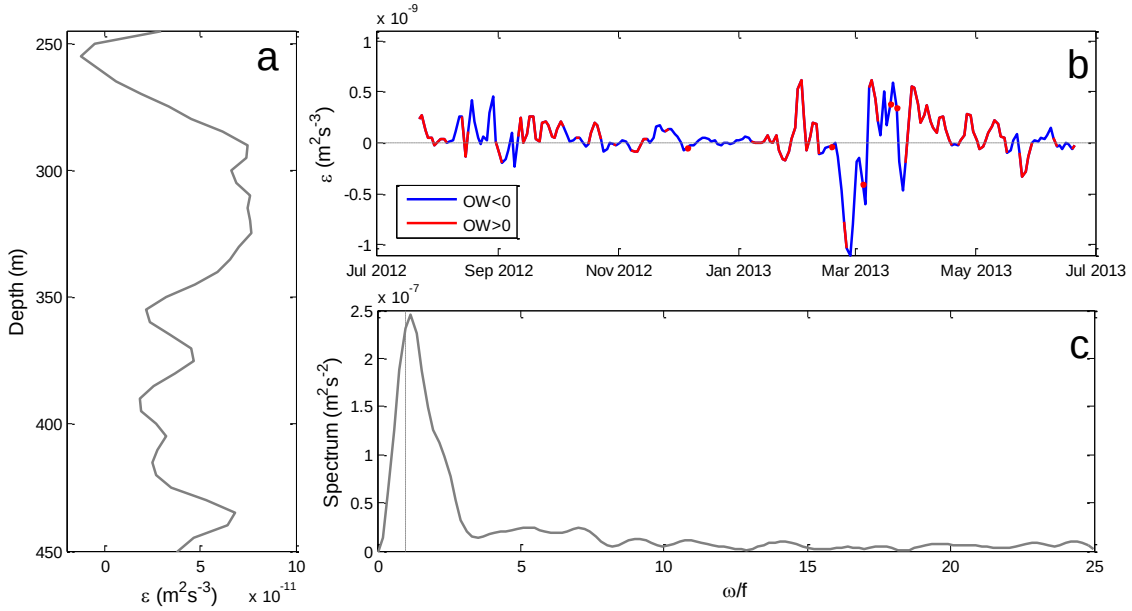

Figure 3.3 (a) The time-mean energy transfer rate ε; (b) The time series of ε averaged within 245-450 m; and (c) Partition of ε in the frequency domain.

The simulation starts from January 1, 2012 and ends on June 30, 2013. Only the results during July 1, 2012-June 30, 2013 are used for analysis, which is coincident with the period of mooring measurements. The initial and boundary conditions are obtained

from Hybrid Coordinate Ocean Model (hycom) data assimilation (Chassignet 2007). The atmospheric forcing is derived from the ERA-Interim reanalysis dataset (Dee et al. 2011) with a horizontal resolution of 0.75^{o} and a temporal resolution of 6 hour.

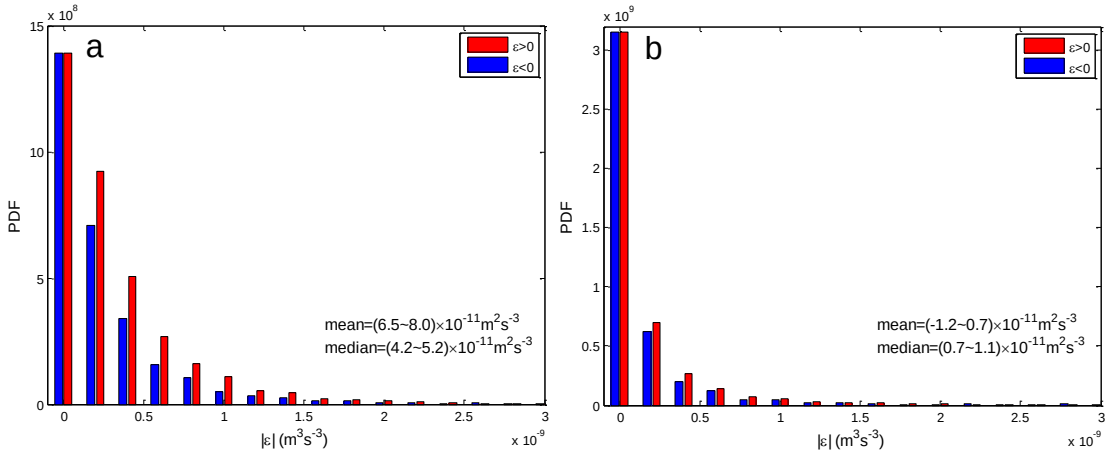

Figure 3.4 The PDFs of ε during (a) positive and (b) negative Okubo-Weiss parameter conditions. The numbers in brackets are the 90%-confidence intervals computed from the bootstrap method.

3.3 Mooring Data Analysis

3.3.1 *Energy exchange between mesoscale and internal waves*

Figure 3.3a displays the vertical distribution of time mean ε. The value of ε is positive throughout 245-450 m (Figure 3.3a). The vertical mean ε is 4.1×10^{-11}m^{2}s^{-3}, corresponding to a permanent energy transfer from mesoscale eddies to internal waves. The time series of ε reveals that the instantaneous energy transfer rate can be both positive and negative (Figure 3.3b). A permanent energy transfer from mesoscale eddies to internal waves appears only to occur in the positive OW parameter condition. This is indicated by the difference of the probability density functions (PDFs) of ε between

positive and negative OW parameter conditions (Figure 3.4). When the OW parameter is positive, the value of PDF with positive ε is significantly larger than that with negative ε, corresponding to a positive energy transfer rate. The mean ε in this case is 7.2×10^{-11}m^{2}s^{-3} and its 90%-confidence interval computed from the bootstrap method is $6.5-8.0\times10^{-11}$ m^{2}s^{-3}. However, the PDF is almost symmetric about $\varepsilon = 0$ when the OW parameter is negative. The mean ε in this case is -0.2×10^{-11} m^{2}s^{-3} and is not significantly different from zero. Therefore, a permanent energy transfer only occurs when the strain dominates relative vorticity.

NIWs make a dominant contribution to the energy transfer (Figure 3.3c). The important role of NIWs is partially due to their energetic wave stress as indicated by the frequency spectrum of horizontal velocity (Figure 3.5). It should be noted that inertial oscillations do not interact with mesoscale eddies as both the normal stress $<uu>-<vv>$ and shear stress $<uv>$ are exactly zero. It is the small departure of wave frequency from f that leads to the energy transfer. Applying the polarization relations of internal waves yields:

$$P_{uu} - P_{vv} = \frac{k^2 - l^2}{k^2 + l^2}\frac{\omega^2 - f^2}{\omega^2 + f^2}(P_{uu} + P_{vv}) \tag{3.7}$$

$$\mathrm{Re}\{P_{uv}\} = \frac{kl}{k^2 + l^2}\frac{\omega^2 - f^2}{\omega^2 + f^2}(P_{uu} + P_{vv}) \tag{3.8}$$

Therefore, the frequency spectra of wave stress are given by the frequency spectrum of horizontal velocity, $P_{uu} + P_{vv}$, multiplied by a factor of $(\omega^2 - f^2)/(\omega^2 + f^2)$. Due to the pronounced peak of $P_{uu} + P_{vv}$ around the inertial frequency, even

$(\omega^2 - f^2)/(\omega^2 + f^2)(P_{uu} + P_{vv})$ exhibits a significant enhancement in the near-inertial band (Figure 3.5), implying the important role of NIWs in the energy exchange between mesoscale eddies and internal waves.

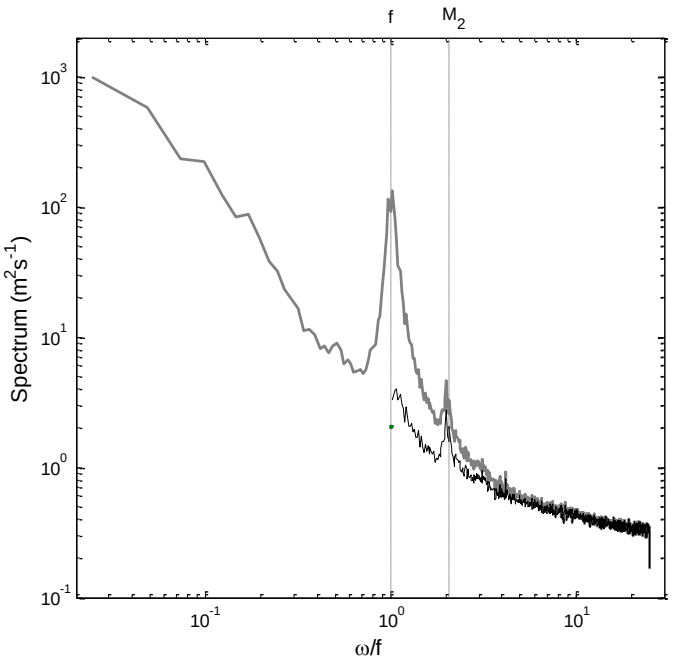

Figure 3.5 The frequency spectrum of horizontal velocity (grey thick) and that multiplied by the factor $(\omega^2 - f^2)/(\omega^2 + f^2)$ (black thin).

NIWs are primarily generated by the wind work on near-inertial motions in the surface mixed layer (Wunsch and Ferrari 2004). Furthermore, as the moorings are close to the critical latitude 28.9°N for M2 tide, the subharmonic parametric instability (PSI) may also play a role in generating NIWs (Zhang et al. 2010). In both cases, the initial azimuth of generated NIWs is expected to be independent from the velocity gradient of mesoscale eddies. Therefore, there would be no net energy exchange between NIWs and

mesoscale eddies unless mesoscale eddies have a permanent influence on the azimuth of NIWs. In the local limit approximation where the relaxation effect dominates, negative correlation of $<uu>-<vv>$ to S_n and $<uv>$ to S_s always holds. In this case, there should be a net energy transfer from mesoscale eddies to NIWs no matter whether the OW parameter is positive or not (Müller 1976). This is, however, not supported by the observations (Figure 3.4). Furthermore, the regression between $<uu>-<vv>$ and S_n suggests that there is no significant negative correlation (Figure 3.6a) when the OW parameter is negative. The estimated horizontal viscosity, $\nu_h = -S_n/(<uu>-<vv>)$ is not significantly different from zero (Figure 3.6a) using a 90%-confidence interval between -7 m^2s^{-1} and 7 m^2s^{-1}. Similar is the case for the regression between $-2<uv>$ and S_s (Figure 3.6b). All these are not consistent with the predictions from the local limit approximation (Müller 1976). It should be noted that the contradiction cannot be reconciled by changing the value of relaxation time scale that is subject to great uncertainties.

However, a net energy transfer only occurring when strain dominates relative vorticity is consistent with the wave capture scenario. In the following discussion, we attempt to evaluate the relevance of wave capture to the extent that the available data permit.

3.3.2 *Role of wave capture in the energy transfer*

In the wave capture scenario, the azimuth θ of horizontal wave vector asymptotically approaches $\theta_a = \arctan(-(U_x+\lambda)/V_x)$ or $\theta_a + \pi$ where λ is half the

square root of OW parameter (See Appendix B for details). In those directions, internal wave extract energy from mesoscale eddies. We introduce θ_w, defined as the minimum between $|\theta_a - \theta|$ and $|\theta_a + \pi - \theta|$, to measure the role of wave capture in energy exchange. When the wave capture mechanism has a negligible influence, θ_w should be uniformly distributed within 0°-90°. But its mean value would be significantly smaller than 45° when the wave capture mechanism indeed plays a role. Figure 3.7 shows the mean θ_w when the OW parameter is positive. While the mean θ_w for high-frequency internal waves is close to 45°, NIWs are associated with a significant reduction of mean θ_w, implying that NIWs are no longer isentropic due to the wave capture mechanism. This is consistent with the dominant contribution of NIWs to the energy transfer rate.

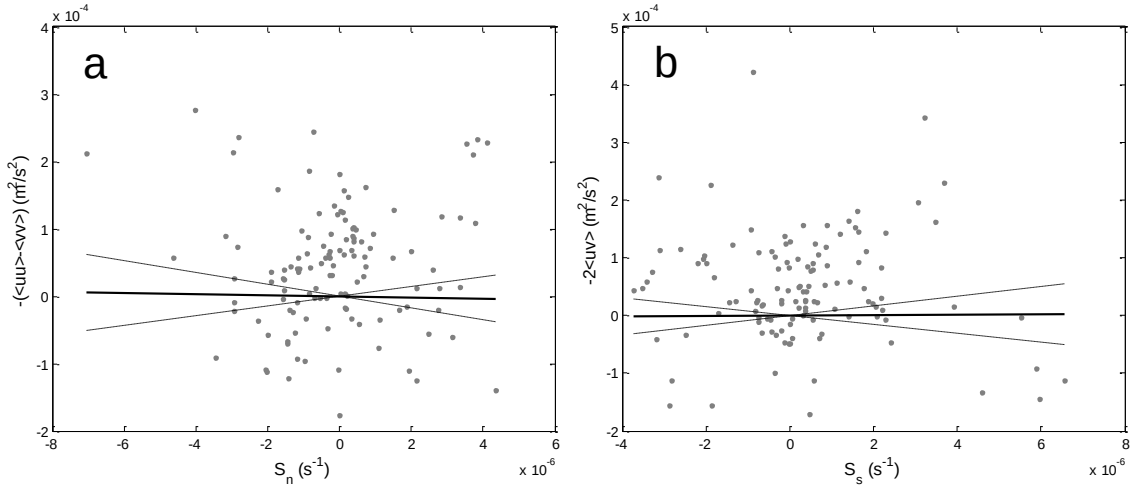

Figure 3.6 (a) The scatterplot of vertical mean $-(<uu>-<vv>)$ versus vertical mean S_n in the period with negative OW parameter. The solid line is the linear regression with its 90%-confidence interval (grey dashed). (b) Similar to (a) but for the scatterplot of $-2<uv>$ versus S_s.

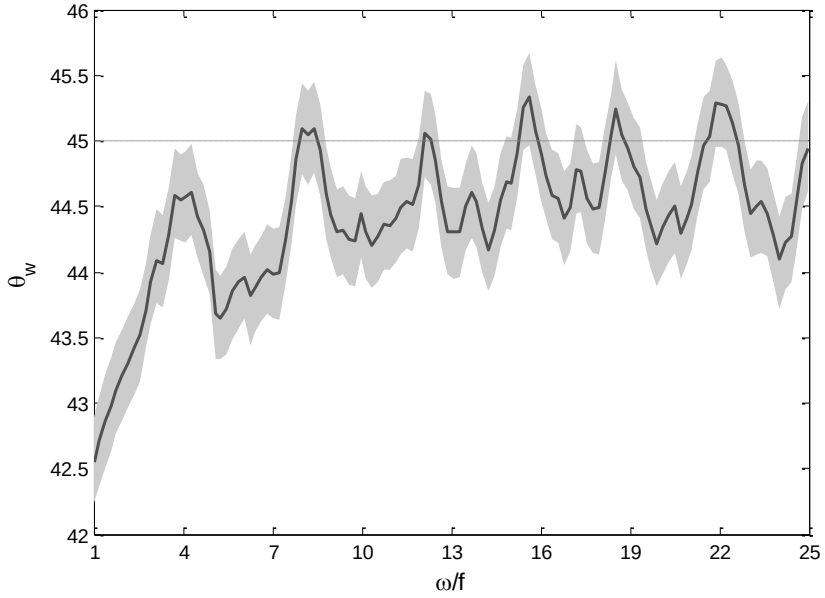

Figure 3.7 The mean θ_w and its 90%-confidence interval computed from the bootstrap method for different wave frequencies.

The mean θ_w for NIWs is only about 3^o smaller than 45^o although the reduction is statistically significant, implying that the wave capture does not always work efficiently when the OW parameter is positive. In the wave capture theory, velocity gradient of mesoscale eddies is assumed to be constant along the ray. However, in the reality it varies both spatially and temporally, leading to variation of θ_a along the ray. If θ_a changes too rapidly, internal waves may not have sufficient time to adjust to the change of θ_a . In this case, it is difficult for θ_w to reach its asymptotic value, i.e., 0^o. Furthermore, the wave capture theory assumes that the mesoscale eddies satisfy the geostrophic relation. Otherwise, a positive OW parameter does not guarantee the occurrence of wave capture. The influence of ageostrophy can be measured by the

difference of λ computed as $\sqrt{S_n^2 + S_s^2 - \varsigma^2}/2$ and λ' computed as $\sqrt{V_x U_y - U_x V_y}$. They are identical when mesoscale eddies are subject to the geostrophic balance but can differ significantly in the presence of strong ageostrophy. Figure 3.8 shows time series of λ and λ'. In most time, they differ by a small amount, suggesting that mesoscale eddies are approximately geostrophic. But pronounced difference also exists. As mentioned above, the wave capture mechanism is expected to work efficiently only when λ and λ' vary slowly in both the space and time domain[2] and are almost identical to each other. We identify three periods (referred to as quasi-steady periods henceforth) satisfying the above requirements by visually inspecting Figure 3.8. In the quasi-steady periods, the energy transfer from mesoscale eddies to internal waves is significantly enhanced (Figure 3.9). The mean energy transfer rate increases to 17.5×10^{-11}m^{2}s^{-3}, more than four times of that during the entire period. In particular, while the length of quasi-stead period is less than18% of the entire period, it accounts for more than 75% of the net energy transfer.

During the quasi-steady period, PDF value of θ_w for NIWs decreases moronically (Figure 3.10a) as θ_w increases, suggesting a highly anisentropic wave field. This provides strong evidence that the wave capture plays a dominant role in the energy exchange between mesoscale eddies and NIWs. To further verify that such a PDF

[2] The spatial variation of velocity gradient cannot be directly estimated due to the limited data. However, a slow variation of velocity gradient in the time domain implies a slow variation in the space domain because the mesoscale eddies migrates westwards.

distribution can be resulted from the wave capture, ray tracing experiments are performed based on the equations (3.1)-(3.3).

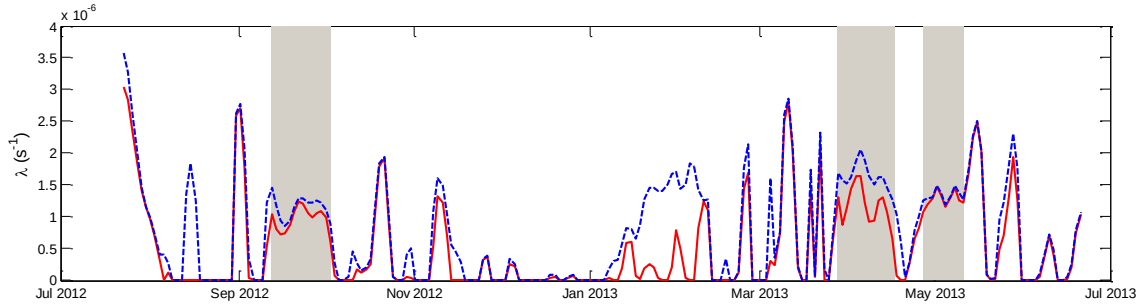

Figure 3.8 Time series of vertical mean λ (red solid) and λ' (blue dashed) within 245-450 m. The shaded regions denote the quasi-stead period.

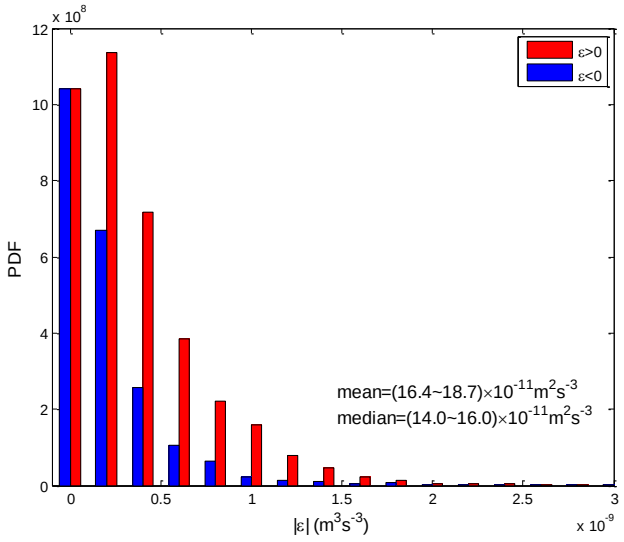

Figure 3.9 The PDF of ε during the quasi-steady periods.

As NIWs are primarily generated by surface wind forcing (Wunsch and Ferrari 2004), the test NIW packets are initially injected on 245 m (the shallowest depth sampled by the ADCPs) and radiate downwards. The initial frequency is chosen

randomly between $f - 2f$ using the observed frequency spectrum of internal wave horizontal velocity as its PDF (Figure 3.5) following Henyey et al. (1986). Similarly, the initial vertical wavenumber is generated using the observed vertical wavenumber spectrum of near-inertial current as its PDF. Then the initial k_h can be uniquely determined based on the dispersion relation while the initial azimuth is assumed to satisfy the uniform distribution within 0°-360°.

The stratification is computed from the CTD profiles collected during the retrieve of moorings. But using a constant stratification rather than the observed one does not make any substantial impact on the simulated PDF of θ_w, implying that the variation of stratification along the ray does not have significant influences on the test waves. The geostrophic currents between 245-450 m are modelled as

$$U = U_T(x, y)F(z) \tag{3.9}$$

$$V = V_T(x, y)F(z) \tag{3.10}$$

where U_T and V_T represent the horizontal structure of geostrophic flow and $F(z)$ represents the vertical structure. $F(z)$ is set to be a linear function decreasing from unity at 245 m to 0.7 at 450 m, which is consistent with the observations derived from the ADCPs (Figure 3.2). U_T and V_T are derived from the sea level anomaly using the geostrophic relation:

$$U_T = -\alpha \frac{g}{f} \eta_y \tag{3.11}$$

$$V_s = \alpha \frac{g}{f} \eta_x \tag{3.12}$$

where η is the sea level anomaly, g is the gravity acceleration, and α is a scaling factor accounting for the difference of geostrophic flow intensity between the sea surface and 245 m. Here α is chosen in such a way that the variance of U_T and V_T is the same to the observed value at 245 m.

Figure 3.11 displays a snapshot of the surface OW parameter computed from sea level anomaly obtained from the Ssalto/Duacs multimission altimeter products of AVISO. Given that the altimeters might not work reasonably well in the coastal region, the initial injection position for test waves is chosen to mimic the situation in the quasi-steady periods rather than the exact location of moorings. This leads to an initial injection position around 92^{o}W, 28^{o}N (referred to as the Site A henceforth). It is characterized by a λ of $\sim 1\times10^{-6}\,\mathrm{s}^{-1}$ with weak spatial variation (The value of λ varies within a factor 2 in a $1^\circ\times1^\circ$ box centered at the site A). As mesoscale eddies propagate westwards at a speed of about 0.1 m s^{-1}, λ at Site A remains almost unchanged within O(10 day). All these features are consistent with the situation in the quasi-steady periods (Figure 3.8). In the ray tracing experiment, 200 test waves are generated. Sensitivity testes suggest that the PDF of θ_w does not vary substantially when the number of test waves increases to 100 and greater. Therefore, 200 test waves should be sufficient to compute the statistics. For each test wave, ray tracing is terminated when the wave (1) propagates into regions associated with negative OW parameter, (2) breaks according to the criterion $|m| > 0.1\,\mathrm{cpm}$, or (3) radiates out of the lower bound, i.e., 450 m. A

normally distributed random noise with a standard deviation of 15^o is added to account for the measurement errors in the observations.

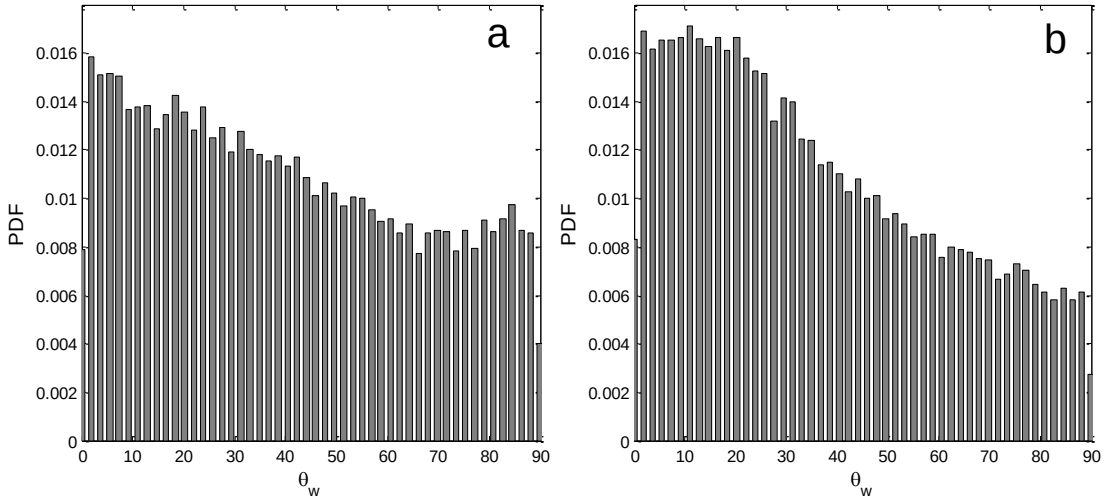

Figure 3.10 The PDF of θ_w (a) during the quasi-stead periods and (b) simulated by the ray tracing experiment.

The simulated PDF of θ_w is qualitatively consistent with the observed one (Figure 3.10a and b). Both exhibit a decreasing value as θ_w increases, implying a tendency of θ to approach θ_a or $\theta_a + \pi$. It should be noted that the relaxation effects have been filtered out in the ray tracing model. Furthermore, absorption of test waves around the critical layer does not occur in the simulation as the vertical shear of geostrophic flow is weak. Finally, the test waves are always located in the strain-dominated region so that the trapping mechanism proposed by Kunze (1985) is unlikely to play an important role. In this model, the wave capture is the only mechanism that can have a permanent influence on θ. Therefore, the similarity between the simulated PDF of θ_w and the

observed one provides further evidence that the wave capture mechanism plays a dominant role in the energy exchange between mesoscale eddies and NIWs during the quasi-steady period.

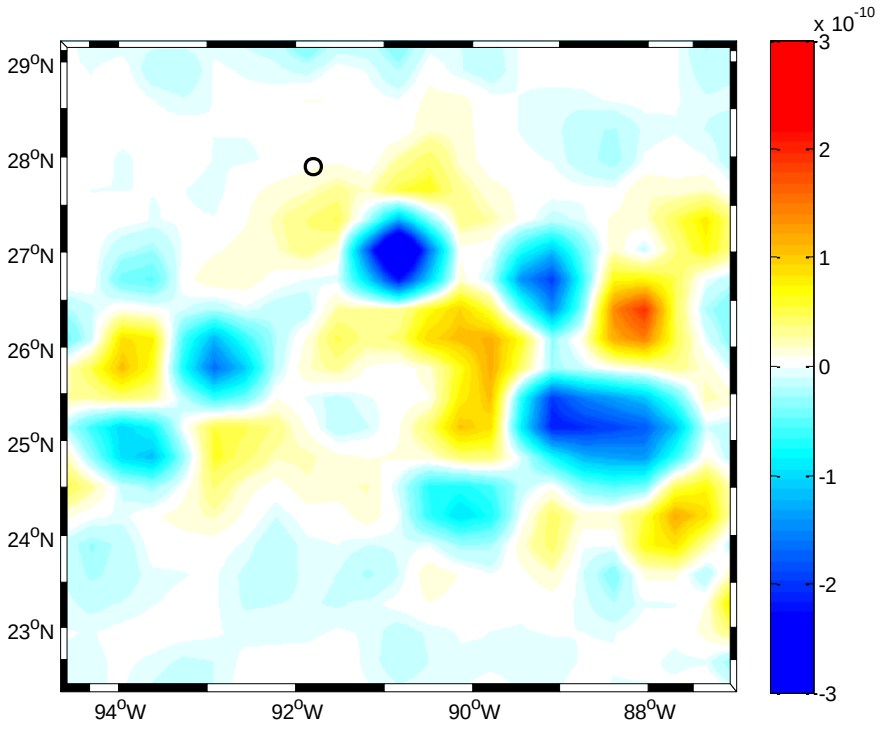

Figure 3.11 A snapshot (September 22, 2012) of OW parameter derived from the sea level anomaly using the geostrophic relation. The sea level anomaly is obtained from the Ssalto/Duacs multimission altimeter products of AVISO.

3.4 Results in High-Resolution Numerical Simulation

In this section, we analyze energy exchange between NIWs and mesoscale eddies in the Gulf of Mexico simulated by the high-resolution ROMS described in Section 3.2.5. Figure 3.12a displays the standard deviation of daily mean sea level anomalies simulated by ROMS. Its overall spatial pattern is consistent with the satellite observations (Figure

3.12b). Both exhibit enhanced activities in the central Gulf of Mexico due to the shedding of mesoscale eddies from the loop current. The magnitude of simulated standard deviation of sea level anomalies is slightly larger than the observed one. But they agree with each other in general within a factor of 2. This gives us confidence that the mesoscale eddies simulated by ROMS are qualitatively reliable.

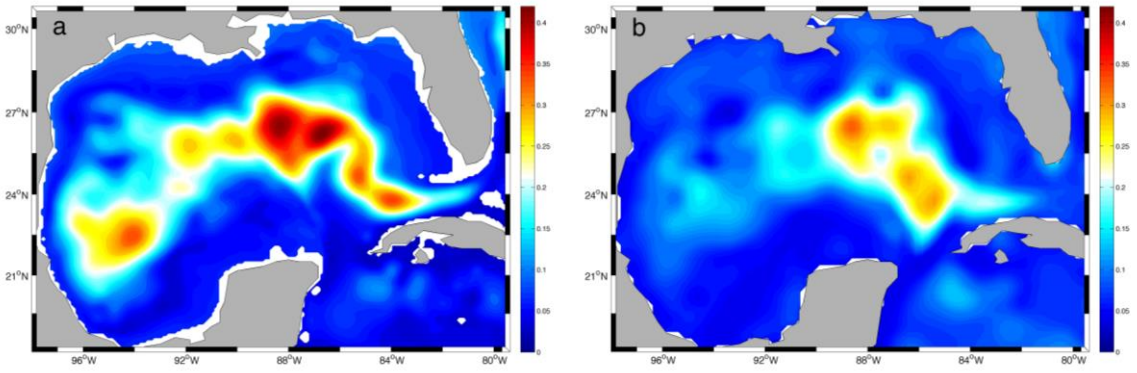

Figure 3.12 The standard deviation of daily mean sea level anomaly during July 2012-Jun 2013 (a) simulated by ROMS and (b) derived from the Ssalto/Duacs multimission altimeter products of AVISO. The unit here is m.

Table 3.1 compares the observed and simulated ε around the mooring location. In the numerical simulation, the mean ε between 245-450 m within 88°W-89°W, 28°N-29°N is 4.2×10^{-11} m^2s^{-3}, close to the value 4.0×10^{-11} m^2s^{-3} derived from the mooring measurements. Furthermore, the simulated ε is close to zero when the OW parameter is negative, suggesting that a permanent energy transfer only occurs under the positive OW parameter condition. This is also consistent with mooring observations. These agreements give us confidence that the energy exchange between mesoscale eddies and NIWs in the numerical simulation is reliable.

Figure 3.13a shows the simulated spatial map of time-mean energy transfer rate $\bar{\varepsilon}$ averaged within 100-1000 m. The top 100-m water column is excluded here because NIWs in the mixed layer and upper thermocline are not free-propagating waves but are directly forced by wind stress. The mechanisms responsible for the energy exchange between mesoscale eddies and wind-forced NIWs are distant from the mechanisms responsible for free-propagating NIWs (Whitt and Thomas 2015) and are beyond the scope of this study. $\bar{\varepsilon}$ is in general positive in the basin of Gulf of Mexico, corresponding to a permanent energy transfer from mesoscale eddies to NIWs. To quantify the role of wave capture in the energy exchange, we compare the contribution to $\bar{\varepsilon}$ during the periods with positive and negative OW parameters, respectively. Here the contribution is measured as:

$$\bar{\varepsilon}_P = \int \varepsilon H(OW) dt / T \tag{3.13}$$

$$\bar{\varepsilon}_N = \int \varepsilon H(-OW) dt / T \tag{3.14}$$

where T represents the length of integration period and $H(x)$ is the Heaviside function. It is obvious from (3.13) and (3.14) that $\bar{\varepsilon} = \bar{\varepsilon}_P + \bar{\varepsilon}_N$. Therefore, $\bar{\varepsilon}_P$ and $\bar{\varepsilon}_N$ can be treated as decompositions of $\bar{\varepsilon}$ into the periods with positive and negative OW parameters.

Similar to $\bar{\varepsilon}$, $\bar{\varepsilon}_P$ is in general positive in the basin of Gulf of Mexico (Figure 3.13b). The basin-mean value of $\bar{\varepsilon}_P$ is 4.3×10^{-11} m^2s^{-3}, about 74% of the basin-mean value 5.8×10^{-11} m^2s^{-3} for $\bar{\varepsilon}$. In particular, the spatial distributions of $\bar{\varepsilon}_P$ and $\bar{\varepsilon}$ agree well with each other. Their squared spatial correlation coefficient R^2 reaches up to 0.95.

In contrast, the value of $\bar{\varepsilon}_N$ is close to zero in most of the regions (Figure 3.13c). R^2 between $\bar{\varepsilon}_N$ and $\bar{\varepsilon}$ is only about 0.12. These facts lend supports to the dominant role of the wave capture mechanism in the energy transfer from mesoscale eddies to NIWs in the Gulf of Mexico.

	Entire Period	OW>0	OW<0
Observations	4.0	7.2	-0.2
Simulations	4.2	5.5	-0.2

Table 3.1 The mean ε between 245-450 m derived from the mooring measurement and numerical simulation within 88°W-89°W, 28°N-29°N during the entire time period, the period with positive OW parameter, and the period with negative OW parameter. The unit here 10^{-11} m^2s^{-3}.

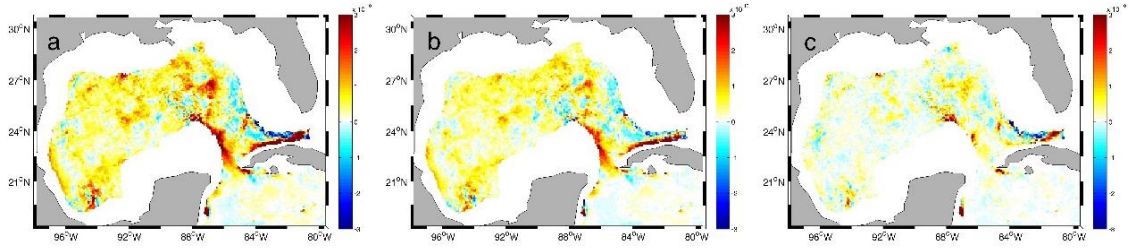

Figure 3.13 The spatial distribution of (a) $\bar{\varepsilon}$, (b) $\bar{\varepsilon}_P$, and (c) $\bar{\varepsilon}_N$ in the high-resolution numerical simulation.

3.5 Summary and Discussion

Long-term mooring array and high-resolution numerical simulations in the Gulf of Mexico are used to analyze the energy exchange between internal waves and mesoscale eddies. Important findings of this study are summarized as follows.

(1) When the OW parameter is negative, the observed energy exchange between mesoscale eddies and internal waves is almost zero. However, there is an evident energy transfer rate of $O(10^{-10} m^2 s^{-3})$ from mesoscale eddies to internal waves when the OW parameter is positive. NIWs play a dominant role in the energy exchange. This is due to both their energetic wave stress and their strong interaction with mesoscale eddies.

(2) The wave capture mechanism plays a key role in the interaction between NIWs and mesoscale eddies. NIWs become highly anisotropic when the OW parameter is positive. The observed PDF of propagation direction of NIWs is consistent with the predictions from the wave capture theory.

(3) The simulated energy exchange between mesoscale eddies and NIWs in the high-resolution numerical simulation is qualitatively consistent with the mooring observations. A permanent energy transfer from mesoscale eddies to NIWs mainly occurs when the OW parameter is positive, lending further supports to the dominant role of wave capture mechanism in the energy exchange between mesoscale eddies and NIWs in the Gulf of Mexico.

It is generally believed that winds provide a primary energy source for oceanic NIWs (Wunsch and Ferrari). However, as mentioned in the introduction, part of the wind work on near-inertial motions is inevitably dissipated within the mixed layer. The wind-generated near-inertial energy flux into the deep ocean is relatively weak. It might be possible that the energy exchange between mesoscale eddies and NIWs provides another important energy source for NIWs in the deep ocean. This conjecture is

supported by our high-resolution numerical simulation. The simulated basin-mean near-inertial wind work is about 0.80 mW m^{-2}. Assuming that about 30% of the near-inertial wind work radiates downwards into the region below 100 m (Alford et al. 2012; Zhai et al. 2009; Furuichi et al. 2008), the winds could provide an energy flux of 0.24 mW m^{-2} for NIWs below 100 m. Meanwhile, the vertically integrated (100-1000 m) energy transfer rate from mesoscale eddies to NIWs is equal to 0.053 mW m^{-2}. This number is about 22% of the wind-induced near-inertial energy flux (0.24 mW m^{-2}) into the region below 100 m, suggesting that the energy exchange between mesoscale eddies and NIWs may play an important role in furnishing NIWs in the deep ocean.

CHAPTER IV

INTERACTIONS BETWEEN NEAR-INERTIAL INTERNAL WAVES AND SUBMESOSCALE INTERNAL WAVE FRONTS

4.1 Introduction

Vertical motion in the ocean plays a key role in exchange of heat, dissolved gasses, and nutrients between the upper and deep ocean. Previous studies indicated that pumping fluxes induced by mesoscale eddies do not provide sufficient nutrient fluxes to sustain observed levels of productivity in subtropical gyres (Oschlies 2002; Martin and Pondaven 2003). Submesoscale phenomena O(<10 km) have received growing attention over the past decade because of their intense vertical motions, due to the ageostrophic nature compared to mesoscale flows. It has been noted that submesoscale motions may play a critical role in vertical transport processes in the ocean. Although supporting evidence for this notion from direct observations has been difficult to come by, modelling studies have shown significantly enhanced vertical transport in high-resolution simulations that explicitly resolve submesoscale processes (e.g., Levy et al. 2001; Mahadevan and Archer 2000; Hansen and Samuelsen 2009; Zhong and Bracco 2013).

Near the surface, energetic submesoscale vertical motions often emerge alongside strong density fronts. Several mechanisms have been proposed for the development of submesoscale motions in the presence of strong horizontal density gradients with O(1) Rossby number, including strain-driven frontogenesis (Hoskins and Bretherton 1972) ,

mixed layer instability (Fox-Kemper et al. 2006), ageostrophic anticyclonic instability (Molemaker et al. 2005), and nonlinear Ekman transport effects (Stern 1965; Thomas and Lee 2005). As density fronts typically become much weaker with increasing depth, submesoscale motions generated by the near-surface mechanisms are mainly confined to surface mixed layer and upper thermocline.

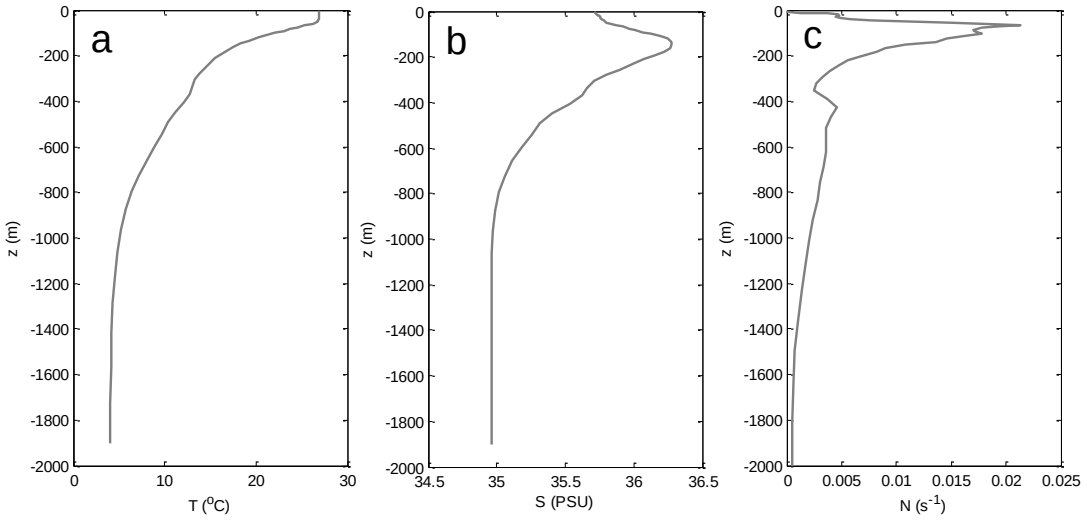

Figure 4.1 The initial profiles in IHS for (a) potential temperature, (b) salinity, and (c) buoyancy frequency.

Zhong and Bracco (2013) recently reported that energetic submesoscale fronts (SMFs) with strong vertical motions do not only occur in the upper ocean but also in the deep Gulf of Mexico. The simulated submesoscale motions at depth in their model are characterized by the first few vertical modes with frequencies significantly higher than the inertial frequency *f*, but their appearance seems nevertheless to be closely associated with the energetic near-inertial internal waves (NIWs) rather than the near-surface

density fronts. This connection between the SMFs and NIWs raises the possibility of a dynamic linkage between them. However, the underlying mechanism remains to be unexplored, which is the focus of chapter.

This chapter is organized as follows. An idealized high-resolution numerical simulation designed to illustrate the relationship between SMFs and NIWs is first presented in Section 4.2. Theoretical solutions governing SMF – NIW interaction are developed in Section 4.3 and are validated against the idealized numerical simulation. A further validation of the theoretical results using a realistic simulation of circulation in the Gulf of Mexico is made in Section 4.4. Conclusions and discussion are finally given in Section 4.5.

4.2 An Idealized High-Resolution Numerical Simulation

4.2.1 *Model description and experiment design*

To explore mechanisms for SMF – NIW interaction, an idealized high-resolution numerical experiment is performed using the Regional Ocean Modelling System (ROMS), which is a free surface 3-D primitive equation model based on hydrostatic and Boussinesq approximations (Shchepetkin and McWilliams 2005). The model is configured over a $20° \times 20°$ domain with a uniform depth of 2000 m. 50 vertical layers are used with 19 layers concentrated in the upper 100 m. The horizontal grid size is set at 1 km $\times$ 1 km, so that submesoscale flows can be explicitly resolved. The nonlocal K-profile parameterization (Large et al. 1994) is used to parameterize vertical mixing, but no horizontal eddy viscosity and diffusivity are used since at 1-km resolution the model

should explicitly resolve mesoscale and submesoscale eddies. The horizontal diffusivity and viscosity are thus taken as their molecular values. Finally, a radiation boundary condition is used at lateral boundaries of the model domain.

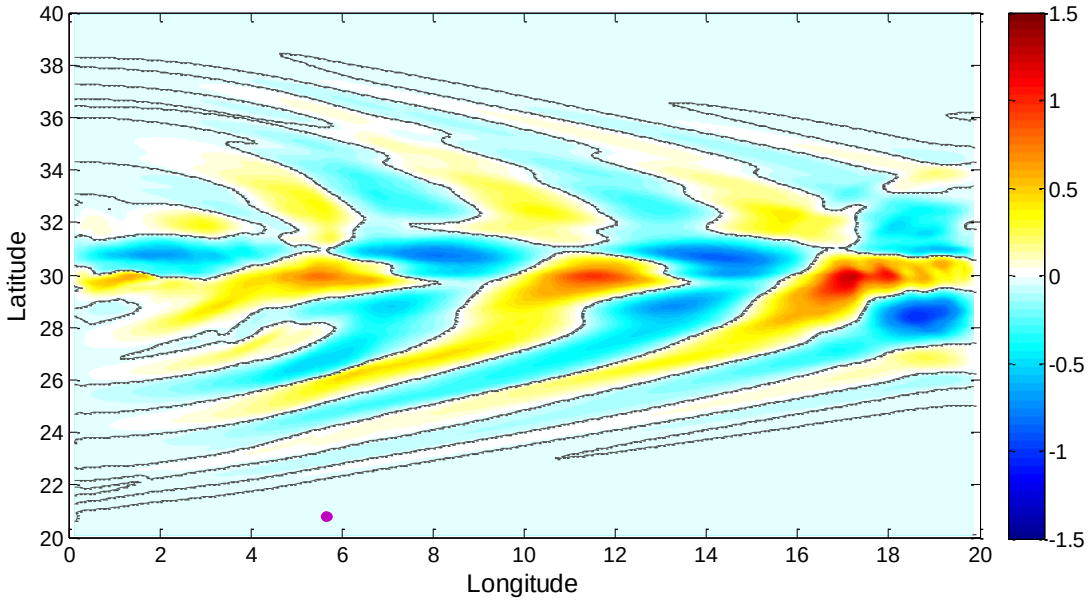

Figure 4.2 Large-scale (>30 km) surface zonal velocity in color (m s^{-1}) on 6d0h in IHS. The grey contours denote the zero (phase) contours of NIWs.

It should be noted that the hydrostatic approximation made by ROMS tends to overestimate the vertical velocity of submesoscale flow due to its much reduced horizontal/vertical aspect ratio (Vitousek and Fringer 2011). This tendency is counter-balanced to some extent by numerical dispersion that mimics the missing physical dispersion due to nonhydrostacy. In fact, previous numerical studies indicated that the numerical dispersion can be tuned to replicate the nonhydrostatic dispersion not resolved

in a hydrostatic model (Shuto 1991; Burwell et al. 2007). Therefore, the hydrostatic approximation is acceptable for our qualitative analysis here.

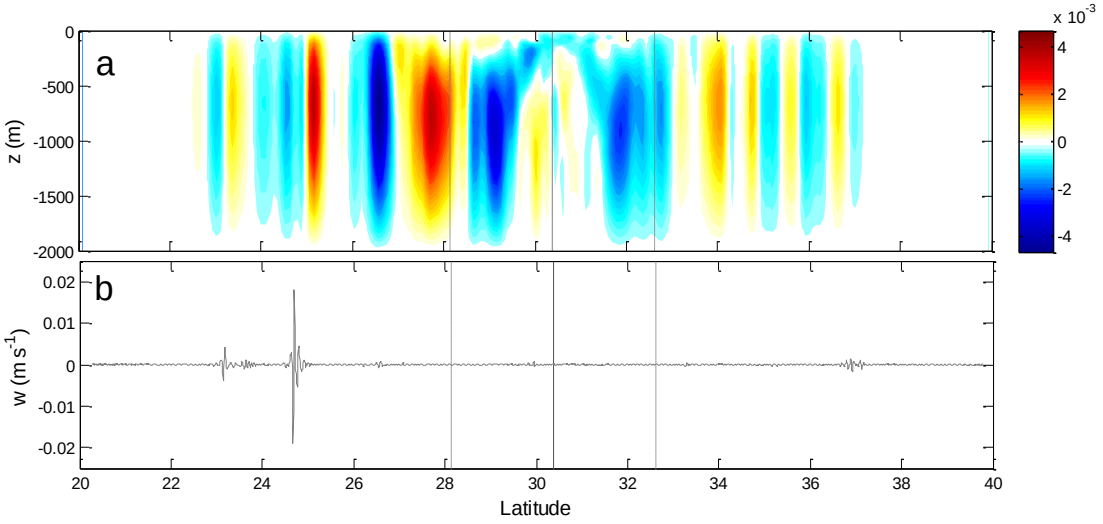

Figure 4.3 (a) Meridional section of large-scale (>30 km) vertical velocity in color (m s^{-1}); (b) Snapshot of submesoscale (<15 km) vertical velocity at 100 m. The values shown here are on 6d0h along the longitude 7.4º in IHS. The solid vertical line denotes the center of the hurricane while the dashed lines mark the edges of the hurricane.

The simulation starts with a quiescent and horizontally homogeneous ocean, in which the temperature and salinity profiles are representative of those in the Gulf of Mexico during summer (Figure 4.1a and b). This stratification is characterized by a sharp thermocline roughly at 65 m with a maximal buoyancy frequency of 2.1×10^{-2} rad s^{-1} (Figure 4.1c). To excite energetic NIWs, an idealized west-to-east moving hurricane at a constant translation speed U_T of 7 m s^{-1} is applied at the center of the model domain as wind forcing. The center of the hurricane is initially located at the middle of the western boundary and moves to the eastern boundary after 4 days and 1 hour (4d1h). We

refer this idealized hurricane simulation as IHS and examine the influences of hurricane-induced NIWs on submesoscale motions in IHS. The wind field associated with the idealized hurricane is constructed following Price (1983). The wind vector is axisymmetric and rotates cyclonically with maximal tangential velocity occurring at radius R=40 km. The maximal wind speed is set at 52 m s^{-1}, corresponding to a category-5 hurricane. Outside the hurricane, the surface heat flux and freshwater flux are set to zero.

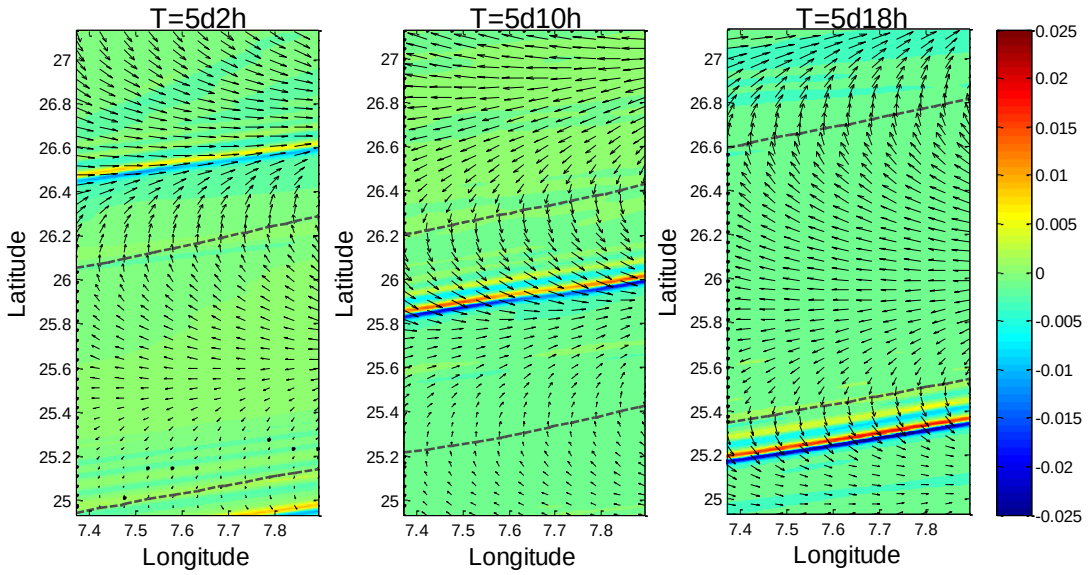

Figure 4.4 Sequential snapshots of vertical velocity (m s^{-1}) at 100 m on 5d2h (left), 5d10h (middle), and 5d18h (right) in IHS. The quivers represent the surface horizontal velocity associated with large-scale (>30 km) NIWs with the grey lines denoting the phase contours.

4.2.2 *IHS result*

In the wake of the idealized hurricane, energetic NIWs are generated as expected from the theory, which are clearly visible in the horizontal sea-surface velocity field

(Figure 4.2). There is a marked cross-track asymmetry for the hurricane-generated NIWs with more energetic waves occurring on the right of the moving storm. This is mainly because the wind stress vector rotates anti-cyclonically on the right of the track and thus is able to resonate with the inertial oscillations there (Price 1981 and 1983). NIW energy emanates away from the hurricane track in different vertical modes. The wave structure far away from track is characterized by the first baroclinic mode (Figure 4.3a), because the first baroclinic mode has the fastest group velocity. The vertical structure near the hurricane track is more complex (Figure 4.3a) because it is composed of many vertical modes. However, even in this case the first baroclinic mode still dominates, implying that most of the wind-induced near-inertial energy is projected onto the first baroclinic mode (Gill 1984).

While NIWs in the hurricane wake are fully expected from the theory and previous studies, a new interesting feature emerging from the high-resolution simulation is the appearance of wave-like SMFs with vigorous vertical velocity of O(0.01 m s^{-1}) at 100 m (Figure 4.3b), which is an order of magnitude larger than that induced by the NIWs at the same depth. These SMFs tend to occur in the far field from the hurricane track and have horizontal scales of O(10 km). As such, models with coarse resolutions (~10 km) cannot capture these SMFs (not shown), consistent with Zhong and Bracco (2013). Similar to the hurricane-generated NIWs, there is also an evident cross-track asymmetry in the SMFs with much more energetic SMFs on the right of the track.

A close examination of the relationship between the simulated NIWs and SMFs reveals that SMFs tend to occur in convergence (downwelling) zones of NIWs (Figure

4.3). To further examine the SMF – NIW relationship, Figure 4.4 shows zoom-in snapshots following the strongest SMF, from which one can clearly see that the SMF is aligned parallel with the phase contours of NIWs and propagates at a speed comparable to the gravity wave speed (~2 m s^{-1}) of the first baroclinic mode. During its propagation in the convergence (downwelling) zone of NIWs, the spatial structure of the SMF does not change significantly while its magnitude grows rapidly in the initial phase before it becomes saturated (Figure 4.4).

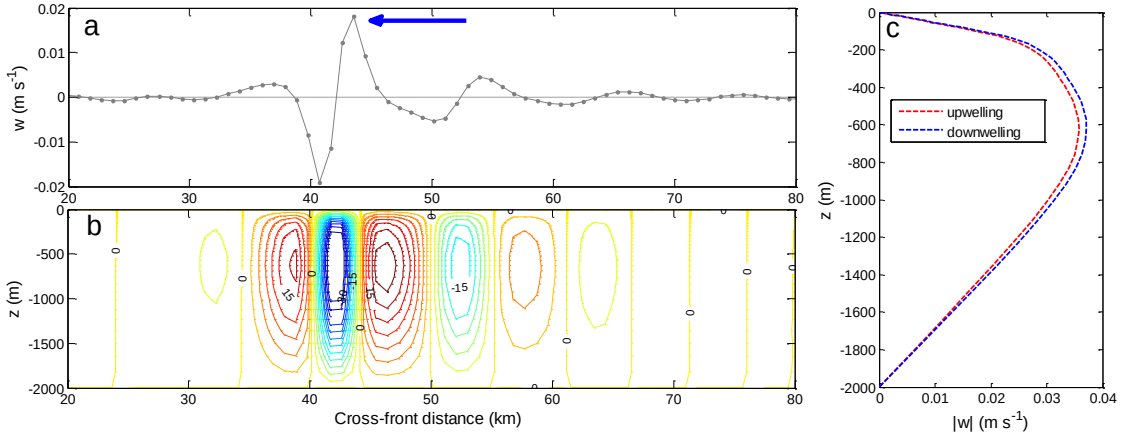

Figure 4.5 (a) Cross-front distribution of vertical velocity at 100 m on 6d0h. The blue arrow indicates the propagation direction. (b) Similar to (a) but for the stream function (m^2s^{-1}). The contour interval here is 3 m^2s^{-1}. (c) shows vertical profiles of vertical velocity magnitude at the center of the upwelling and downwelling cells.

As a first order approximation, these wave-like SMFs can be characterized by two-dimensional waves with a zero wavenumber in along-frontal direction (Figure 4.4). As such, a 2-D stream function Ψ in the space of cross-front distance and depth can be used to represent SMF structure. Figure 4.5 shows the stream function of the strongest

SMF on 6d0h. It is evident from Figure 4.5 that the SMF's wavelength is on the order of 10 km and its vertical structure is well represented by the first baroclinic mode. A frequency spectral analysis indicates that its frequency is much higher than the inertial frequency and its wavenumber – frequency relationship agrees well with the short internal gravity wave dispersion relationship, $\omega = c_1 k$, where c_1 is the gravity wave speed of the first baroclinic mode (Figure 4.6).

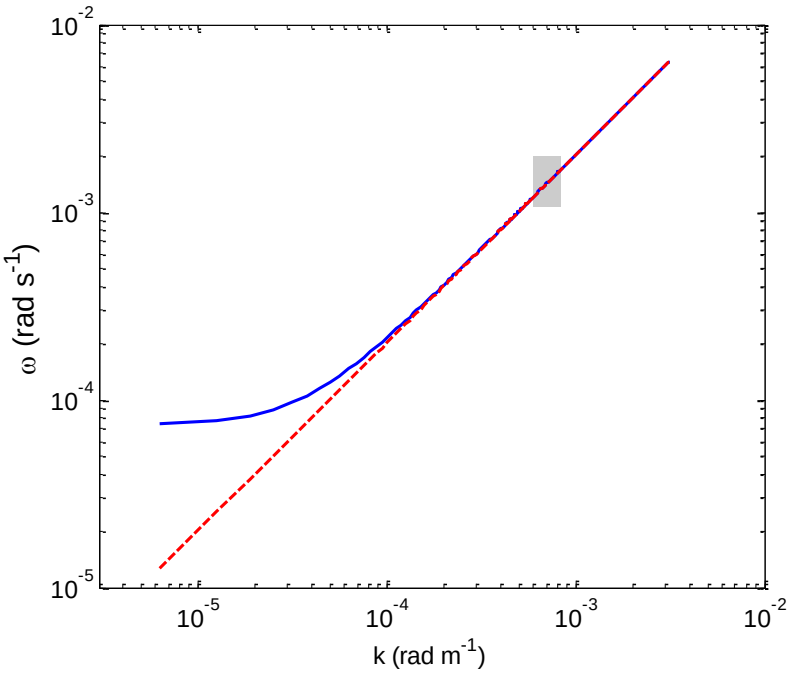

Figure 4.6 The dispersion relation curve for dispersive waves ($\omega^2 = c_1^2 k^2 + f^2$; blue solid) and non-dispersive waves ($\omega^2 = c_1^2 k^2$; red dashed). The grey shaded region denotes the estimated frequency and horizontal wavenumber range of the SMF during 4d14h-6d0h.

In summary, the idealized numerical simulation reveals that NIWs can support energetic SMFs in their convergence (downwelling) zones. The SMFs possess strong

vertical velocity of O(0.01 m s^{-1}) and typical wavelength of O(10 km), propagating in the same direction as NIWs at a speed comparable to the gravity wave speed of the first baroclinic mode. Its vertical structure projects strongly onto the first baroclinic mode. The wavenumber – frequency relationship of SMFs is consistent with the dispersion relationship for short internal gravity waves. It is, thus, highly suggestive that the SMFs belong to the non-dispersive high-frequency internal gravity wave family and their generation is through interaction with NIWs.

4.3 Analytical Analysis

In this section, we present analytical analyses to understand the dynamic mechanism governing the SMF – NIW interaction revealed by IHS described above. The fact that both the SMFs and NIWs project strongly onto the first baroclinic mode allows us to reduce the complexity of the problem by considering only one vertical mode. In the following, we first present analysis in a reduced gravity model framework represented by the first vertical baroclinic mode and then in a more generalized dynamic model framework.

4.3.1 *SMF – NIW interaction in a reduced gravity model*

As shown in IHS, the wave-like SMFs are characterized by moderate vertical motions when they first emerge (Figure 4.4a) and then amplify rapidly (Figure 4.4b and c) through interactions with NIWs. During the initial development phase, SMFs' nonlinearity is expected to play a minor role. The governing equations for the SMFs at this stage in a reduced gravity model framework can be linearized as:

$$\frac{Du}{Dt} = fv - g'\frac{\partial \eta}{\partial x} - u\frac{\partial U}{\partial x} - v\frac{\partial U}{\partial y} \tag{4.1a}$$

$$\frac{Dv}{Dt} = -fu - g'\frac{\partial \eta}{\partial y} - u\frac{\partial V}{\partial x} - v\frac{\partial V}{\partial y} \tag{4.1b}$$

$$\frac{D\eta}{Dt} + \frac{\partial (H+Z)u}{\partial x} + \frac{\partial (H+Z)v}{\partial y} = -\eta(\frac{\partial U}{\partial x} + \frac{\partial V}{\partial y}) \tag{4.1c}$$

where g' is the reduced gravity, H is the undisturbed layer thickness, U (u), V (v), and Z (η) are the horizontal currents and interface elevation produced by NIWs (SMFs). Here we use g'=0.01 m s^{-2} and H =400 m so that the gravity wave speed of the reduced gravity model matches with that of the first baroclinic mode in IHS.

The β effect is unlikely to play a major role here due to the small horizontal scale of SMFs. So an *f*-plane approximation is applied. Furthermore, the system is isotropic horizontally under the *f*-plane approximation, so that the Cartesian coordinate can be rotated to make the y-axis parallel with the submesoscale front. As both the SMFs and NIWs exhibit much less variations in the along-front direction than in the cross-front direction (Figure 4.4), the system at hand can be reduced to a unidirectional wave equation by setting all the *y*-derivatives to zero. Note that the negligence of *y*-derivatives does not lose generality for the following derivations. The conclusions in this section are also valid in the more generalized case where the submesoscale front does not necessarily align parallel with the phase contours of NIWs (See Appendix C for details).

To proceed systematically, it is essential to introduce non-dimensional parameters. Let L_N and U^* represent the horizontal scale and horizontal velocity magnitude of

NIWs and L_S and $u*$ for SMFs. A reasonable measurement of a wave's horizontal scale should be the reciprocal of horizontal wavenumber. In IHS, NIWs away from the hurricane's track are associated with wave frequencies between $1.1-1.8f$. According to the dispersion relation of NIWs, their horizontal wavenumbers range from $0.5L_D^{-1}-1.5L_D^{-1}$, where L_D is the deformation radius. Therefore, it is reasonable to choose $L_N = L_D$. We then use these characteristic scales to non-dimensionalize both dependent and independent variables, denoted by primes, as follows:

$$\eta = \frac{c}{g'} u * \eta' \tag{4.2a}$$

$$Z = \frac{c}{g'} U * Z' \tag{4.2b}$$

$$(u, v) = u * (u', v') \tag{4.2c}$$

$$(U, V) = U * (U', V') \tag{4.2d}$$

$$x = L_S x' \tag{4.2e}$$

$$t = \frac{L_S}{c} t' \tag{4.2f}$$

where $c = \sqrt{g'H}$ =2 m s^{-1} is the gravity wave speed. Note that *x* is scaled based on SMFs, so that any *x*-derivative term related to NIWs should be multiplied by a scaling factor $\varepsilon = L_S / L_D$ to account for the difference in the horizontal scales between NIWs and SMFs. Substituting (4.2) into (4.1) yields the following non-dimensionalized equations:

$$\frac{\partial u}{\partial t} + \delta U \frac{\partial u}{\partial x} = \varepsilon v - \frac{\partial \eta}{\partial x} - \delta \varepsilon u \frac{\partial U}{\partial \varepsilon x} \tag{4.3a}$$

$$\frac{\partial v}{\partial t}+\delta U\frac{\partial v}{\partial x}=-\varepsilon u-\delta\varepsilon u\frac{\partial V}{\partial \varepsilon x} \tag{4.3b}$$

$$\frac{\partial \eta}{\partial t}+\delta U\frac{\partial \eta}{\partial x}+(1+\delta Z)\frac{\partial u}{\partial x}+\delta\varepsilon u\frac{\partial Z}{\partial \varepsilon x}=-\delta\varepsilon\eta\frac{\partial U}{\partial \varepsilon x} \tag{4.3c}$$

where $\delta = U^*/c$ is a non-dimensional parameter measuring the nonlinearity of NIWs. For the sake of neatness, we drop the primes for the non-dimensional variables here and hereinafter. The non-dimensional energy equation associated with (4.3) is given by

$$\frac{\partial e}{\partial t}+\delta\frac{\partial Ue}{\partial x}=-\frac{\partial F}{\partial x}+P_b+P_s \tag{4.4}$$

where $e=(1+\delta Z)(u^2+v^2)/2+\eta^2/2$ is the energy density of SMFs, $F=(1+\delta Z)u\eta$ the energy flux of SMFs due to pressure work, and P_b (P_s) the potential (kinetic) energy exchange between NIWs and SMFs:

$$P_b=-\frac{1}{2}\delta\varepsilon\eta^2\frac{\partial U}{\partial \varepsilon x}$$

$$P_s=-\delta\varepsilon(1+\delta Z)(uu\frac{\partial U}{\partial \varepsilon x}+uv\frac{\partial V}{\partial \varepsilon x})$$

δ and ε are two small (<<1) non-dimensional parameters in (4.3) and their relatively magnitude can be measured as: $\chi=\delta/\varepsilon$. Here we assume that χ is of order of unity. In fact, χ ranges from 0.5 to 2 for U^*=0.05~0.2 m s^{-1}, c=2 m s^{-1}, $L_S=10/(2\pi)$ km, and $f=6\times10^{-5}$ s^{-1}. As the temporal and horizontal scales of SMFs and NIWs are well separated, a multiple-scale method (Johnson 2005) is applied to obtain the solutions to (4.3). Introduce the fast and slow variables

$$\frac{\partial\varphi}{\partial t}=-\omega,\ \frac{\partial\varphi}{\partial x}=k,\ T_1=\varepsilon t,\ T_2=\varepsilon^2 t,\ X=\varepsilon x \tag{4.5}$$

where ω and k are the local wave frequency and wavenumber of SMFs, assuming that the SMFs can be approximated as a plane wave locally, which is indeed the case in IHS (Figure 4.5). The temporal and spatial derivatives can then be expanded, in virtue of the chain rule, as:

$$\frac{\partial}{\partial t} = -\omega \frac{\partial}{\partial \varphi} + \varepsilon \frac{\partial}{\partial T_1} + \varepsilon^2 \frac{\partial}{\partial T_2} \tag{4.6a}$$

$$\frac{\partial}{\partial x} = k \frac{\partial}{\partial \varphi} + \varepsilon \frac{\partial}{\partial X} \tag{4.6b}$$

Expand the SMFs in asymptotic series of ε:

$$u = \sum_{n=0}^{\infty} \varepsilon^n u_n \tag{4.7a}$$

$$v = \sum_{n=0}^{\infty} \varepsilon^n v_n \tag{4.7b}$$

$$\eta = \sum_{n=0}^{\infty} \varepsilon^n \eta_n \tag{4.7c}$$

$$\omega = \sum_{n=0}^{\infty} \varepsilon^n \omega_n \tag{4.7d}$$

Insert (4.6) and (4.7) into (4.3) and the O(1) terms yield

$$-\omega_0 \frac{\partial u_0}{\partial \varphi} = -k \frac{\partial \eta_0}{\partial \varphi} \tag{4.8a}$$

$$-\omega_0 \frac{\partial v_0}{\partial \varphi} = 0 \tag{4.8b}$$

$$-\omega_0 \frac{\partial \eta_0}{\partial \varphi} + k \frac{\partial u_0}{\partial \varphi} = 0 \tag{4.8c}$$

The solutions to (4.8) are

$$(u_0, v_0, \eta_0) = \mathrm{Re}\{A_0 e^{i\varphi}(\frac{k}{\omega_0}, 0, 1)\} \tag{4.9}$$

with the non-dimensional dispersion relation $\omega_0^2 = k^2$. This dispersion relation corresponds to non-dispersive high-frequency internal gravity waves and is consistent with the analysis of the SMFs in IHS (Figure 4.6). Without loss of any generality, we choose $\omega_0 = k$.

The equations at O(ε) are

$$-\omega_0 \frac{\partial u_1}{\partial \varphi} + k\frac{\partial \eta_1}{\partial \varphi} = -\frac{\partial u_0}{\partial T_1} + \omega_1 \frac{\partial u_0}{\partial \varphi} - \frac{\partial \eta_0}{\partial X} - \chi k U \frac{\partial u_0}{\partial \varphi} \tag{4.10a}$$

$$-\omega_0 \frac{\partial v_1}{\partial \varphi} = -u_0 \tag{4.10b}$$

$$-\omega_0 \frac{\partial \eta_1}{\partial \varphi} + k\frac{\partial u_1}{\partial \varphi} = -\frac{\partial \eta_0}{\partial T_1} + \omega_1 \frac{\partial \eta_0}{\partial \varphi} - \frac{\partial u_0}{\partial X} - \chi k U \frac{\partial \eta_0}{\partial \varphi} - \chi k Z \frac{\partial u_0}{\partial \varphi} \tag{4.10c}$$

Removing singularity requires that the following equations must be satisfied:

$$\frac{\partial A_0}{\partial T_1} + \frac{\partial c_g^0 A_0}{\partial X} = 0 \tag{4.11a}$$

$$\omega_1 = \chi k(U + \frac{1}{2}Z) \tag{4.11b}$$

where $c_g^0 = \partial\omega_0 / \partial k = 1$ is the group velocity of SMFs at O(1).

In virtue of (4.11), the solutions to (4.10) are

$$u_1 = \mathrm{Re}\{(A_1 - \frac{\chi}{2} Z A_0) e^{i\varphi}\} \tag{4.12a}$$

$$v_1 = \mathrm{Re}\{-\frac{i}{k} A_0 e^{i\varphi}\} \tag{4.12b}$$

$$\eta_1 = \mathrm{Re}\{A_1 e^{i\varphi}\} \tag{4.12c}$$

Incorporating η_1 into η_0, then the equation for SMF energy density can be derived by substituting (4.9) and (4.12) into (4.4). Averaging over one wavelength or wave period yields:

$$(\frac{\partial < e >}{\partial T_1} + \varepsilon \frac{\partial < e >}{\partial T_2} + c_g \frac{\partial < e >}{\partial X}) = \delta\sigma < e > + O(\varepsilon^2) \tag{4.13}$$

where $< e >= \mid A_0 \mid^2 / 2 + O(\varepsilon^2)$ is the energy density of SMFs averaged over one wavelength or wave period and

$$c_g = \frac{\partial \omega_0 + \omega_1}{\partial k} = (1 + \frac{\delta}{2} Z) + \delta U \tag{4.14a}$$

$$\sigma = -(\frac{5}{2}\frac{\partial U}{\partial X} + \frac{1}{2}\frac{\partial Z}{\partial X}) \tag{4.14b}$$

c_g is the group velocity that determines the energy propagation of SMFs. The term δU represents the Doppler shift by the horizontal velocity of NIWs, while the term $\delta Z / 2$ represents the refraction effect due to the layer thickness fluctuations induced by NIWs. The coefficient σ is the growth rate determining the rate of change of $< e >$ in a reference frame following the SMFs.

The validity of (4.13) and (4.14) is demonstrated in Appendix D. Its dynamics can be understood by introducing the wave action density $< n >=< e > / \omega_i$, where

$\omega_i = (1+\delta Z/2)k$ is the intrinsic wave frequency (Bretherton and Garrett 1968). According to the ray tracing relation (Lighthill 1978), we have

$$\frac{1}{\omega_i}\left(\frac{\partial \omega_i}{\partial T_1} + \varepsilon \frac{\partial \omega_i}{\partial T_2} + c_g \frac{\partial \omega_i}{\partial X}\right) = -\delta \frac{3}{2}\frac{\partial U}{\partial X} + O(\varepsilon^2) \tag{4.15}$$

Substitute (4.15) into (4.13) yields

$$\frac{\partial < n >}{\partial T_1} + \varepsilon \frac{\partial < n >}{\partial T_2} + \frac{\partial c_g < n >}{\partial X} = O(\varepsilon^2) \approx 0 \tag{4.16}$$

(4.16) is the wave action conservation equation for SMFs during their interactions with NIWs.

In the aid of (4.16), the dynamics in the modulation of SMFs by NIWs become quite clear. Following the SMFs, $< n >$ is amplified in the convergence (downwelling) zone of NIWs due to the converged wave action flux of SMFs, as indicated by (4.14a) and (4.16). Furthermore, the converging currents of NIWs can cause an increase in *k*, and thus ω_i by squeezing the phase contours of SMFs. These two effects work in sync, resulting in a rapid amplification of $< e >$. It should be noted that the growth of submesoscale vertical velocity would be even more pronounced as $< w^2 >= \omega_i^2 < \eta^2 >= \omega_i^2 < e >$. This explains why the vertical velocity associated with SMFs is rapidly enhanced in the convergence (downwelling) region of NIWs (Figure 4.4). The opposite occurs in the divergence (upwelling) zone of NIWs.

4.3.2 *SMF – NIW interaction in a 3-D primitive equation system*

The above analysis based on the reduced gravity model suggests that the underlying dynamics for the modulation of SMFs by NIWs can be understood in terms of wave

action conservation of SMFs in the presence of NIWs. In this section, we examine whether this mechanism is also applicable to a 3-D primitive equation system.

The linearized equations for the SMFs in the 3-D primitive equation system are:

$$\frac{\partial u}{\partial t}+U\frac{\partial u}{\partial x}+W\frac{\partial u}{\partial z}=fv-\frac{\partial p}{\partial x}-(u\frac{\partial U}{\partial x}+w\frac{\partial U}{\partial z}) \tag{4.17a}$$

$$\frac{\partial v}{\partial t}+U\frac{\partial v}{\partial x}+W\frac{\partial v}{\partial z}=-fu-(u\frac{\partial V}{\partial x}+w\frac{\partial V}{\partial z}) \tag{4.17b}$$

$$0=-\frac{\partial p}{\partial z}+b \tag{4.17c}$$

$$\frac{\partial b}{\partial t}+U\frac{\partial b}{\partial x}+W\frac{\partial b}{\partial z}+wN^2=-(u\frac{\partial B}{\partial x}+w\frac{\partial B}{\partial z}) \tag{4.17d}$$

$$\frac{\partial u}{\partial x}+\frac{\partial w}{\partial z}=0 \tag{4.17e}$$

where $\mathbf{u}$ ($\mathbf{U}$) is the three-dimensional velocity, b (B) the buoyancy, and p (P) the density-normalized pressure associated with the SMFs (NIWs). Similar to the derivations in Section 4.3.1, the coordinate has been rotated to make the y-axis parallel with the submesoscale front so that all the y-derivatives can be dropped.

Approximating the SMFs and NIWs by the first baroclinic mode yields

$$u=u_1(x,y,t)\frac{\sqrt{H}F_1(z)}{c_1},\ U=U_1(x,y,t)\frac{\sqrt{H}F_1(z)}{c_1} \tag{4.18a}$$

$$v=v_1(x,y,t)\frac{\sqrt{H}F_1(z)}{c_1},\ V=V_1(x,y,t)\frac{\sqrt{H}F_1(z)}{c_1} \tag{4.18b}$$

$$w=w_1(x,y,t)\sqrt{H}N_0G_1(z),\ W=W_1(x,y,t)\sqrt{H}N_0G_1(z) \tag{4.18c}$$

$$p = g'\eta_1(x,y,t)\frac{\sqrt{H}F_1(z)}{c_1},\ P = g'Z_1(x,y,t)\frac{\sqrt{H}F_1(z)}{c_1} \tag{4.18d}$$

$$b = g'\eta_1(x,y,t)N^2\frac{\sqrt{H}G_1(z)}{c_1},\ b = g'Z_1(x,y,t)N^2\frac{\sqrt{H}G_1(z)}{c_1} \tag{4.18e}$$

where H is the water depth, c_1 the gravity wave speed of first-baroclinic mode, $N_0 = c_1/H$ the reference buoyancy frequency, $g' = c_1^2/H$ the reduced gravity, and G_1 (F_1) the eigenfunction for the first-baroclinic-mode vertical (horizontal) velocity satisfying $\int_{-H}^{0} G_1 G_1 N^2 dz = 1$ ($\int_{-H}^{0} F_1 F_1 dz = c_1^2$). Here the scaling factors are chosen in such a way that

$$\int_{-H}^{0}\frac{1}{2}(u^2+v^2)dz = \frac{H}{2}(u_1^2+v_1^2) \tag{4.19a}$$

$$\int_{-H}^{0}\frac{b^2}{2N^2}dz = \frac{1}{2}g'\eta_1^2 \tag{4.19b}$$

Substituting (4.18) into (4.17), the governing equations for the first-baroclinic-mode SMFs in presence of first-baroclinic-mode NIWs become:

$$\frac{\partial u_1}{\partial t} + sU_1\frac{\partial u_1}{\partial x} + \frac{s}{2}W_1\frac{u_1}{H} = fv_1 - g'\frac{\partial \eta_1}{\partial x} - s(u_1\frac{\partial U_1}{\partial x} + \frac{1}{2}w_1\frac{U_1}{H}) \tag{4.20a}$$

$$\frac{\partial v_1}{\partial t} + sU_1\frac{\partial v_1}{\partial x} + \frac{s}{2}W_1\frac{v_1}{H} = -fu_1 - s(u_1\frac{\partial V_1}{\partial x} + \frac{1}{2}w_1\frac{V_1}{H}) \tag{4.20b}$$

$$\frac{\partial \eta_1}{\partial t} + \frac{s}{2}U_1\frac{\partial \eta_1}{\partial x} + sW_1\frac{\eta_1}{H} + w_1 = -s(\frac{1}{2}u_1\frac{\partial Z_1}{\partial x} + w_1\frac{Z_1}{H}) \tag{4.20c}$$

$$\frac{\partial u_1}{\partial x} - \frac{w_1}{H} = 0 \tag{4.20d}$$

where *s* is the non-dimensionalized projection coefficient defined as $\sqrt{H}c_1^{-3}\int_{-H}^{0}F_1^3 dz$.

If stratification is uniform, i.e., *N* is a constant, F_1 will be a cosine function. In this case, *s* will be exactly zero, so that SMFs do not interact with NIWs. However, for a realistic oceanic setting where stratification is typically very strong in the upper several hundred meters and becomes much weaker in the deep ocean, the magnitude of F_1 is much stronger in the upper ocean according to the WKB theory, leading to a positive value of *s*. In IHS, *s* is estimated to be 2.2 due to the sharp thermocline centered around 65 m (Figure 4.1c).

Using the same non-dimensional variables as in (4.2), the non-dimensionalized form of (4.20) is:

$$\frac{\partial u_1}{\partial t} + s\delta(U_1 \frac{\partial u_1}{\partial x} + \frac{\varepsilon}{2} W_1 u_1) = \varepsilon v_1 - \frac{\partial \eta_1}{\partial x} - s\delta(\varepsilon u_1 \frac{\partial U_1}{\partial \varepsilon x} + \frac{1}{2} w_1 U_1) \tag{4.21a}$$

$$\frac{\partial v_1}{\partial t} + s\delta(U_1 \frac{\partial v_1}{\partial x} + \frac{\varepsilon}{2} W_1 v_1) = -\varepsilon u_1 - s\delta(\varepsilon u_1 \frac{\partial V_1}{\partial \varepsilon x} + \frac{1}{2} w_1 V_1) \tag{4.21b}$$

$$\frac{\partial \eta_1}{\partial t} + s\delta(\frac{1}{2} U_1 \frac{\partial \eta_1}{\partial x} + \varepsilon W_1 \eta_1) + w_1 = -s\delta(\frac{\varepsilon}{2} u_1 \frac{\partial Z_1}{\partial \varepsilon x} + w_1 Z_1) \tag{4.21c}$$

$$\frac{\partial u_1}{\partial x} - w_1 = 0 \tag{4.21d}$$

For the sake of neatness, we again dropped prime for the non-dimensional variables here and hereinafter.

The associated non-dimensional energy equation for SMFs is:

$$\frac{\partial}{\partial t} e_1 = P_{a,1} + P_{w,1} + P_{s,1}^H + P_{s,1}^V + P_{b,1} \tag{4.22}$$

where

$$e_1 = \frac{1}{2}(u_1^2 + v_1^2 + \eta_1^2)$$

$$P_{a,1} = -s\delta(\frac{\partial}{\partial x}U_1\frac{u_1^2 + v_1^2}{2} + \frac{\partial}{\partial x}\frac{1}{2}U_1\frac{\eta_1^2}{2})$$

$$P_{w,1} = -\frac{\partial}{\partial x}(\eta_1 u_1)$$

$$P_{s,1}^H = -s\delta\varepsilon(u_1 u_1\frac{\partial U_1}{\partial \varepsilon x} + u_1 v_1\frac{\partial V_1}{\partial \varepsilon x})$$

$$P_{s,1}^V = -\frac{s\delta}{2}(u_1 w_1 U_1 + v_1 w_1 V_1)$$

$$P_{b,1} = -\frac{s\delta\varepsilon}{2}u_1\eta_1\frac{\partial Z_1}{\partial \varepsilon x} - s\delta w_1\eta_1 Z_1 - \frac{3s\delta\varepsilon}{2}\frac{\partial U_1}{\partial \varepsilon x}\frac{\eta_1^2}{2}$$

Here e_1 is the energy density of SMFs, $P_{a,1}$ the advection term, $P_{w,1}$ the convergence/divergence of SMF energy flux due to pressure work, $P_{s,1}^H$ ($P_{s,1}^V$) the kinetic energy exchange between NIWs and SMFs through the horizontal (vertical) shear of NIWs, $P_{b,1}$ the potential energy exchange between NIWs and SMFs. Note that $P_{s,1}^V$ has no counterpart in the energy equation (4.4) of the reduced gravity model since the horizontal current is vertically uniform in the reduced gravity model.

Following the procedures in Section 4.3.1 (See Appendix E for details), the SMF energy equation (4.22) at $O(\varepsilon)$ is:

$$\frac{\partial < e_1 >}{\partial T_1} + \varepsilon\frac{\partial}{\partial T_2} < e_1 > + c_g\frac{\partial < e_1 >}{\partial X} = \delta\sigma < e_1 > + O(\varepsilon^2) \qquad (4.23)$$

where

$$c_g = 1 + s\delta\frac{Z_1}{2} + s\delta U_1 \tag{4.24a}$$

$$\sigma = -s(\frac{5}{2}\frac{\partial U_1}{\partial X} + \frac{1}{2}\frac{\partial Z_1}{\partial X} - \frac{1}{2}V_1) \tag{4.24b}$$

c_g is the group velocity of SMFs while σ is the growth rate determining the rate of change of $<e>$ in a reference frame following the SMFs.

The SMF energy equation (4.23) and (4.24) in the 3-D primitive equation system is similar to its counterpart in the reduced gravity model, i.e., (4.13) and (4.14). Aside from the projection coefficient *s*, there is only one difference between the growth rate σ in the primitive equation system (4.24b) and reduced gravity system (4.14b). Compared to (4.14b), an additional term $sV_1/2$ arises in (4.24b). This term corresponds to $P_{s,1}^V$ in (4.22) and originates from the energy exchange between SMFs and NIWs through the vertical shear of NIWs (See Appendix E for details). This energy exchange is absent in the reduced gravity model as there is no vertical shear in the reduced gravity model. However, as demonstrated in the following subsection (see Figure 4.7), the energy exchange due to the vertical shear of NIWs appears to play a negligible role compared to the remaining terms in (4.24b). Therefore, the wave action of SMFs in the 3-D primitive equation system can still be treated as a conserved quantity in practice (See Appendix E for details).

The above analysis suggests that the fundamental conclusions derived from the reduced gravity model hold to a large extent in the 3-D primitive equation system. The modulation of SMFs by NIWs can be primarily understood in terms of wave action

conservation of SMFs in the presence of NIWs. SMFs are enhanced in the convergence (downwelling) region of NIWs and damped in the divergence (upwelling) region.

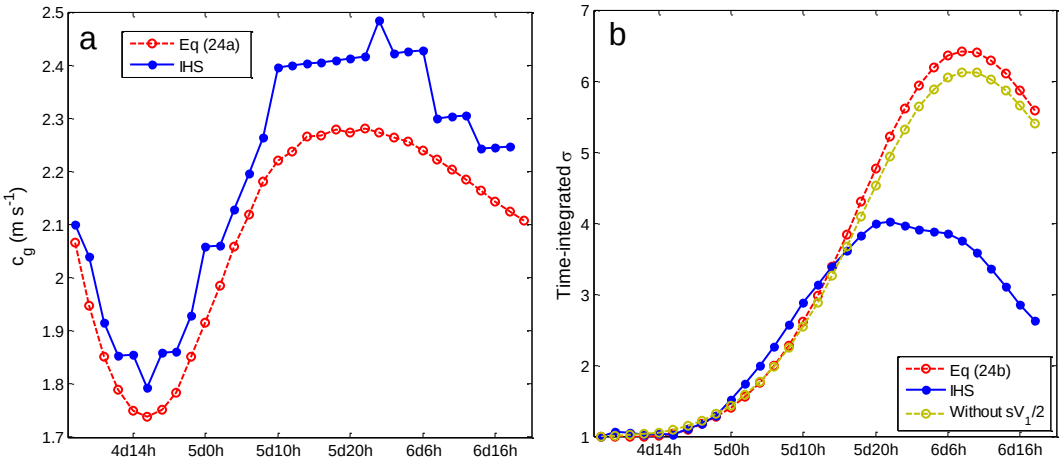

Figure 4.7 (a) Propagation speed of the SMF and (b) evolution of its energy density normalized by the initial value. The solid and dashed lines denote the values computed from the numerical simulation in IHS and analytical solutions, respectively.

4.3.3 *Validation of analytical analysis using IHS*

The validity of (4.23) and (4.24) can be assessed based on the comparisons with the numerical simulation in IHS. Here we track the SMF shown in Figure 4.4 from its initial emergence until it propagates out of the model domain. To separate the SMF from the NIWs, a spatial low-pass filter is applied to isolating the NIWs with a cutoff wavelength of 30 km while the SMF is attained by using a spatial high-pass filter with a cutoff wavelength of 15 km. The group velocity derived from (4.24a) is consistent with that evaluated from IHS (Figure 4.7a). Particularly, they agree well with each other during the initial growth stage of the SMF (before 5d10h) with a discrepancy less than 10%.

The larger error afterwards might be attributed to the nonlinear effects as the amplitude of the SMF has undergone a substantial increase (Figure 4.7b). As u_1 within the submesoscale front is of the same direction to the group velocity (Figure 4.5), the nonlinear advection would contribute to a faster propagation in the numerical experiment than the linear model predicts.

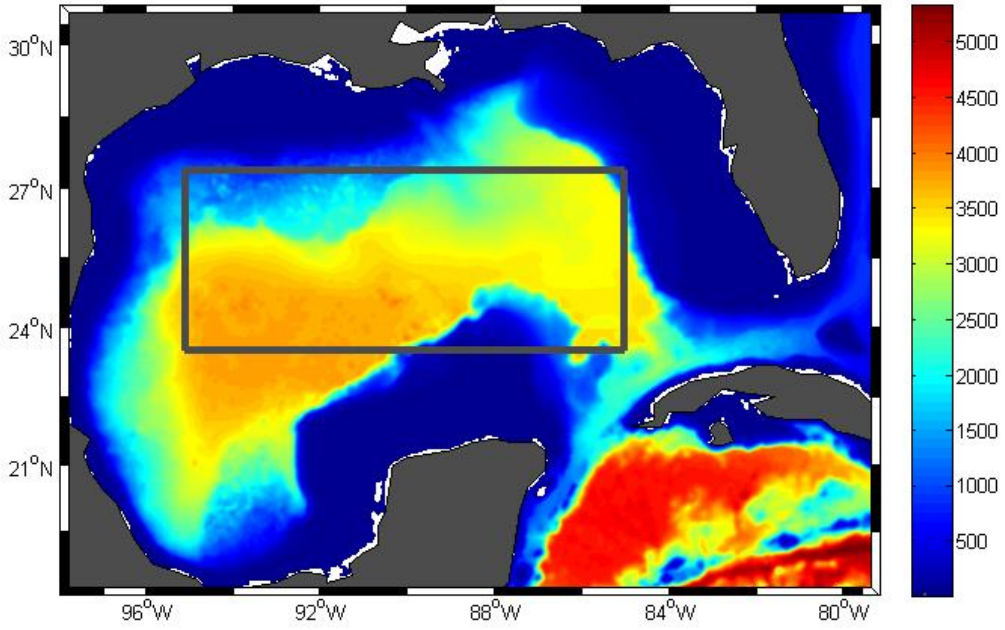

Figure 4.8 Model domain in GoM-C and GoM-F. The contours are the topography. Only the region within the grey box is used for analysis in this study.

The growth rate derived from (4.24b) is also consistent with the numerical simulation in IHS (Figure 4.7b). Here the evolution of the SMF is divided into three stages. The first stage is the growing stage (before 5d10h) when its magnitude increases rapidly. The second is the saturation stage (5d10h-6d6h) with little magnitude variation. The final one is the decay stage (after 6d6h) when the magnitude starts to attenuate. The

theoretical model (4.24b) shows good skills in the growing stage. Furthermore, it qualitatively reproduces the attenuation of the SMF in the decay stage due to the reversed sign of σ. Not surprisingly, (4.24b) overestimates the growth rate during the saturation stage when the nonlinear effects become important.

Finally, it should be noted that the energy exchange due to the vertical shear of NIWs, i.e., $sV_1/2$ in (4.24b), plays a negligible role compared to the remaining terms in (4.24b). σ computed with and without the term $sV_1/2$ (Figure 4.7b) agree well with each other with a discrepancy less than 5%.

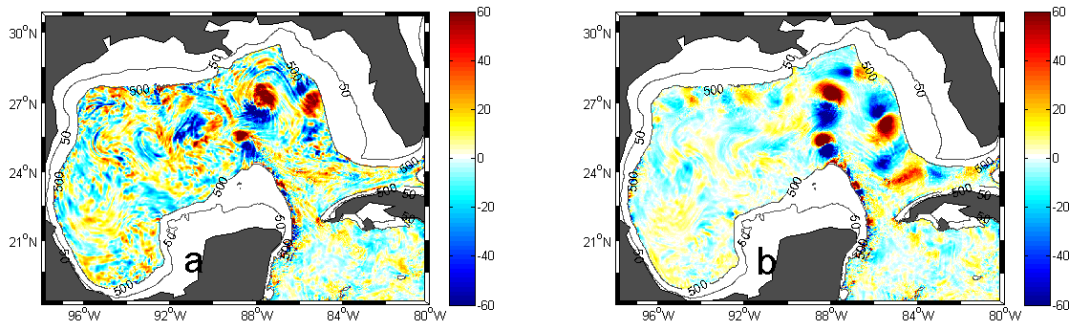

Figure 4.9 Snapshots of vertical velocity at 300 m in (a) GoM-C and (b) GoM-F. The unit here is m/day.

4.4 Validation of Analytical Analysis Using Realistic Numerical Simulations

4.4.1 *Model configurations*

In Section 4.2 and 4.3, we proposed a new mechanism for intense SMFs through SMF – NIW interaction based on the idealized numerical simulation and theoretical analyses. To test whether the theory has any relevance to reality, we perform two more realistic simulations using ROMS configured for the entire Gulf of Mexico and forced

by reanalysis atmospheric forcing (Figure 4.8). The horizontal resolution of ROMS is set at 3 km with 60 vertical layers. The nonlocal K-profile parameterization (Large et al. 1994) is again used to parameterize vertical mixing.

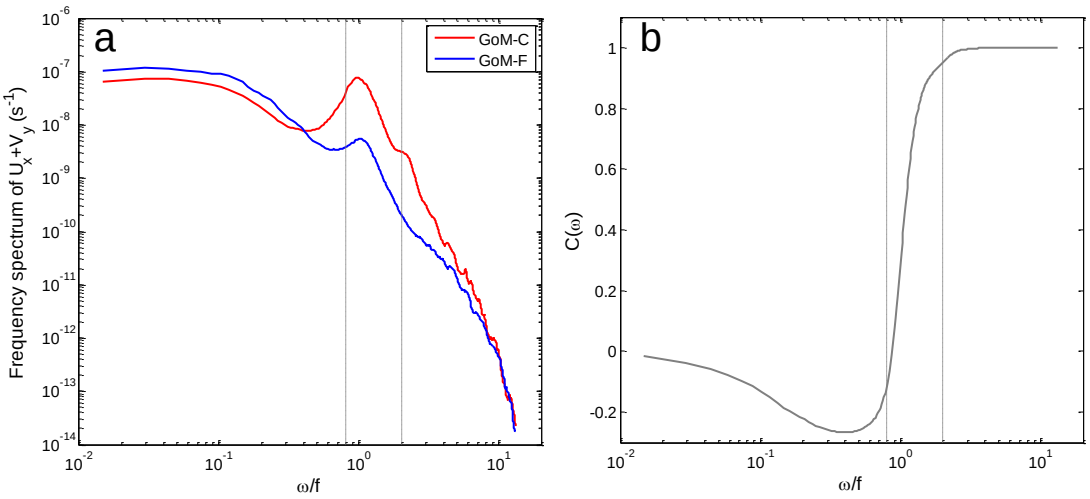

Figure 4.10 (a) The regional mean frequency spectrum of $\nabla\mathbf{U}$ at 100 m in GoM-C (red) and GoM-F (blue). (b) The cumulated contribution $C(\omega)$ to the difference of $\nabla\mathbf{U}$ variance between GoM-C and GoM-F . The grey dashed lines denote the near-inertial (0.8-2 *f*) band.

Both simulations start from March 21, 2010 using Hybrid Coordinate Ocean Model (hycom) data assimilation (Chassignet 2007) as the initial and boundary conditions and last for 90 days. Only the data in the last 20 days are used for the analysis shown below. One simulation, which is referred to as control run (hereinafter GoM-C), is forced by the 6-hourly and 0.25° atmospheric surface variables (e.g., wind and surface air temperature) obtained from the ERA-Interim reanalysis dataset (Dee et al. 2011). The other simulation, which is referred to as the filtered run (hereinafter GoM-F), uses the same atmospheric variables as in GoM-C except that the winds are daily averaged. As the

inertial period in the Gulf of Mexico ranges from 1.6 day at 18°N to 1 day at 30°N, the daily averaged winds contain little variance at inertial and higher frequencies, leading to much suppressed NIWs in GoM-F. Therefore, a comparison between these two experiments with identical model resolutions and physics parameterizations can provide a useful test to our theory.

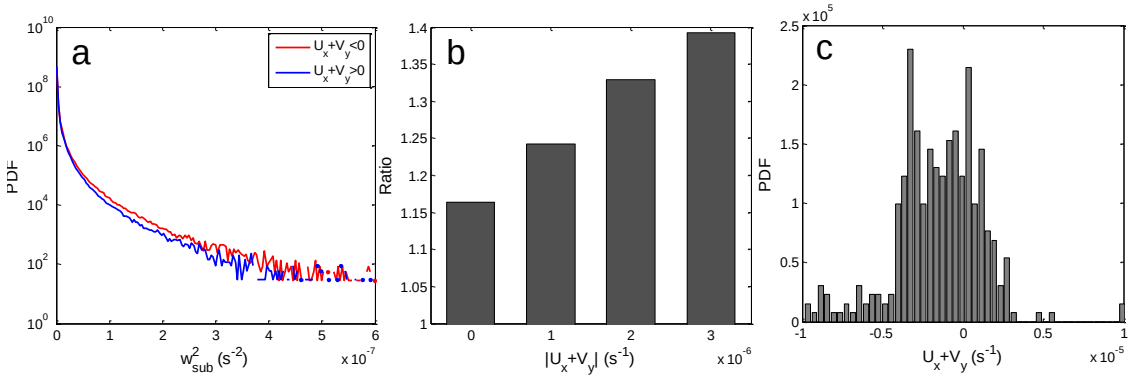

Figure 4.11 (a) PDFs of w^2_{sub} in the convergence region with negative $\nabla\mathbf{U}$ (red) and in the divergence region with positive $\nabla\mathbf{U}$ (blue). (b) The ratio of mean w^2_{sub} in the strong convergence region to that in the strong divergence region with strong convergence/divergence defined as the maginitude of $\nabla\mathbf{U}$ larger than 0, 1×10^{-6}, 2×10^{-6} and 3×10^{-6} s^{-1}, respectively. (c) The conditional PDF of $\nabla\mathbf{U}$ given $|w_{sub}|>50$ m/day.

4.4.2 *Submesoscale vertical motions in GoM-C and GoM-F*

Figure 4.9a shows a snapshot of vertical velocity at 300 m in GoM-C on June 10, 2010. Abundant energetic SMFs with vertical velocity magnitude larger than 50 m/day can be clearly seen. There are also energetic NIWs, as evident by a pronounced peak in the frequency spectrum of large-scale horizontal convergence/divregnce $\nabla\mathbf{U}$ at 100 m (Figure 4.10a). Here the large-scale motions are isolated by a spatial low-pass filter with

a cutoff wavelength of 30 km. More than 60% of $\nabla\mathbf{U}$ varaince comes from the near-inertial (0.8-2*f*) band, suggesting that NIWs make dominant contribution to the large-scale convergenec/divergence. The energetic NIWs are likely linked to the sea breeze whose period is close to the inertial period in the northern Gulf of Mexico (Zhang et al. 2009).

To test the theory developed in Section 4.2 and 4.3 in a realistic setting, we first compare the submesoscale vertical velocity in the convergence ($\nabla\mathbf{U} < 0$) and divergence ($\nabla\mathbf{U} > 0$) zones of large-scale flows in GoM-C. Here the submesoscale motions are isolated using a spatial high-pass filter with a cutoff wavelength of 15 km. $\nabla\mathbf{U}$ is computed using the large-scale horizontal velocity at 100 m. Using velocity at different depths between the surface and 200 m does not change the results presented below significantly. To minimize topographic effects, we applied the analysis to the central Gulf of Mexico (Grey box in Figure 4.8) with water depth larger than 1000 m.

Consistent with the analytical solutions, in the ocean interior (at 300 m) the submesoscale vertical velocity w_{sub} exhibits evident asymmetry between the convergence ($\nabla\mathbf{U} < 0$) and divergence ($\nabla\mathbf{U} > 0$) zones of large-scale flows (Figure 4.11a). The mean w_{sub}^2 in the large-scale flow convergence zone is 1.6×10^{-9} m^2s^{-2}, 16% larger than that in the divergence zone (statistically significant at <1% significance level based on a Wilcoxon ranksum test). This difference is more striking between strong convergence and divergence zones (Figure 4.11b). For instance, w_{sub}^2 is 40% larger in the strong convergence ($\nabla\mathbf{U} < -3.0\times10^{-6}$ s^{-1}) region than in the strong divergence

($\nabla\mathbf{U} > 3.0\times10^{-6}$ s^{-1}) region (statistically significant at <1% significance level). Futhermore, the strong submesoscale vertical motions are found to mainly occur in the convergence region with more than 70% of the strong w_{sub} events of $|w_{sub}|$>50 m/day identified in the convergence ($\nabla\mathbf{U} < 0$) zones (Figure 4.11c). In contrast, only less than 1% of the strong w_{sub} events appear in the strong divergence ($\nabla\mathbf{U} > 3.0\times10^{-6}$ s^{-1}) regions, providing further supports for the theory.

In GoM-F, the NIWs are substantially suppressed as the daily mean winds do not contain any energy in the near-inertial band (Figure 4.10a) for the Gulf of Mexico. Consequently, $\nabla\mathbf{U}$ becomes much weaker in GoM-F than in GoM-C. $\nabla\mathbf{U}$ variance at 100 m decreases from 3.4×10^{-12} s^{-2} in GoM-C to 1.5×10^{-12} s^{-2} in GoM-F. To further quantify the contribution of different frequency bands to the difference of $\nabla\mathbf{U}$ variance between GoM-C and GoM-F, we introduce a cumulated contribution $C(\omega)$ in the frequency domain defined as:

$$C(\omega) = \frac{\int_0^{\omega} P_{con}(s) - P_{fir}(s)ds}{\int_0^{\omega_N} P_{con}(s) - P_{fir}(s)ds} \tag{4.25}$$

where P is the mean frequency spectrum of $\nabla\mathbf{U}$ averaged over the central Gulf of Mexico (Grey box in Figure 4.8), ω_N is the Nyquist frequency, and the subscripts *con* and *fir* represent GoM-C and GoM-F, respectively. $C(2f) - C(0.8f)$ is about 1.05, suggesting that the decrease of $\nabla\mathbf{U}$ variance in GoM-F is mainly due to the much weaker NIWs (Figure 4.10b). It should be noted that the subinertial (<0.8*f*) $\nabla\mathbf{U}$ variance is slighyly weaker in GoM-C than in GoM-F. The reasons remain unclear. Possible

candiates include the dissipation of geostrophic flow due to the horizontal eddy viscosity induced by NIWs (Müller 1976; Polzin 2010).

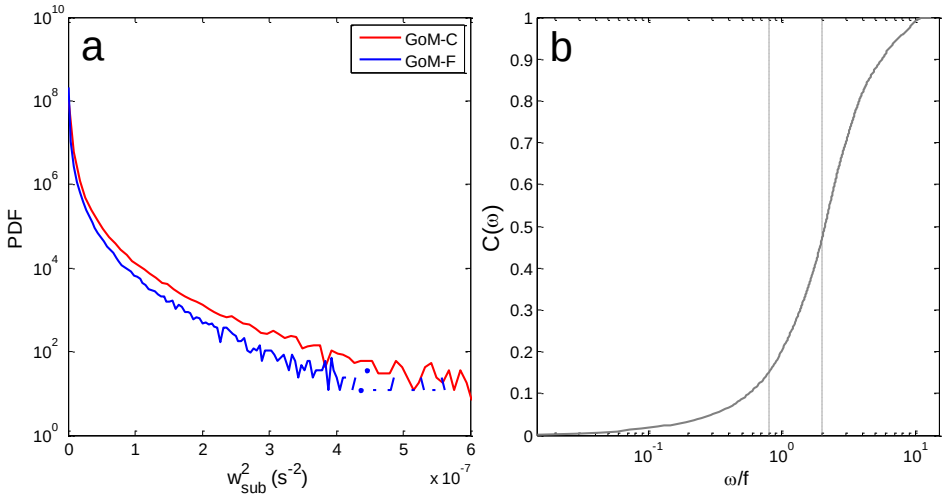

Figure 4.12 (a) PDFs of w^2_{sub} at 300 m in GoM-C (red) and GoM-F (blue); (b) Cumulated contribution $C(\omega)$ to the the difference of w^2_{sub} between GoM-C and GoM-F.

In spite of the fact that both GoM-C and GoM-F are conducted at the same resolution of 3 km, submeoscale vertical motions become significantly weaker in the latter with no energetic SMFs emerging (Figure 4.9b). The mean w^2_{sub} is 1.5×10^{-9} m^2s^{-2} in GoM-C, while it decreases to 0.8×10^{-9} m^2s^{-2} in GoM-F (Figure 4.12a). Such a reduction is statistically significant at <1% significance level, suggesting the important influences of NIWs on the development of submesoscale vertical motions. We note that about 55% of w^2_{sub} difference between GoM-C and GoM-F comes from the superinertial (>2*f*) frequency band while the near-inertial and subinertial bands only account for 31% and 14% of the difference, respectively (Figure 4.12b). Therefore, the bulk of the

submeososcale vertical motions in GoM-C is not directly associated with NIWs but with superinertial motions. As the 6-hourly winds used in GoM-C do not contain any energy in the superinertial frequency band, the energetic submesoscale superinertial motions in GoM-control are unlikely to be directly furnished by winds but are most likely generated through interactions with NIWs. This observation provides futher support for the theoretical finding.

4.5 Summary and Discussion

Submesoscale O(<10 km) motions and their interactions with NIWs are studied theoretically and numerically using high-resolution ROMS simulations. Consistent with the recent modeling study by Zhong and Bracco (2013), SMFs with energetic vertical motions in the ocean interior are found to be closely associated with the NIWs. We present a dynamic mechanism for the SMF development in the presence of background NIWs. It shows that in convergence (downwelling) regions of NIWs, energy flux of the submesoscale motions converges and the energy is transferred from the NIWs to submesoscale motions, leading to enhanced submesoscale vertical velocity. The opposite is true in divergence (upwelling) zones of NIWs. The underlying dynamics can be understood in terms of wave action conservation of submesoscale motions in the presence of background NIWs. This mechanism is analogous to the modulation of surface gravity wave strength by internal waves proposed by Alpers (1985).

The validity of the theoretical finding is confirmed by high-resolution ROMS simulations forced by realistic atmospheric forcing in the Gulf of Mexico. The

simulations show that submesoscale vertical motions within convergence (downwelling) and divergence (upwelling) regions of NIWs exhibit a marked asymmetry with significantly stronger submesoscale vertical velocity occurring in the convergence (downwelling) regions. By removing near-inertial wind forcing, the NIWs are substantially suppressed, which further leads to a significant decrease in the submesoscale vertical motions, suggesting that NIWs can contribute to the development of submesoscale vertical motions.

The proposed SMF-NIW interaction mechanism in this study is different from the strain-driven frontogenesis (Hoskins and Bretherton 1972), mixed layer instability (Fox-Kemper et al. 2006), ageostrophic anticyclonic instability (Molemaker et al. 2005), and nonlinear Ekman transport (Stern 1965; Thomas and Lee 2005) mechanisms which are related to the strong horizontal density gradients and require a O(1) Rossby number. In particular, the enhancement of submesoscale vertical motions by NIWs is not confined to the mixed layer but extends to the ocean interior. This effectively enhances the vertical exchange of heat, dissolved gases, and nutrients between the upper and deep ocean.

CHAPTER V

CONCLUSIONS AND FUTURE WORK

Interactions between near-inertial internal waves (NIWs) and mesoscale-to-submesoscale flows are investigated in this dissertation from three aspects. First, downward radiation of NIWs from mixed layer to deep ocean in idealized baroclinic and barotropic geostrophic flows is theoretically analyzed based on the equation proposed by Young and Ben Jelloul (1997) (the YJB equation). We explore the mechanisms responsible for the dispersion of NIWs in the presence of a baroclinic geostrophic flow and compare the differences between baroclinic and barotropic background flows. Second, energy exchange between NIWs and mesoscale eddies is estimated based on both observations and high-resolution numerical simulation in the Gulf of Mexico. The relative importance of various mechanisms governing NIW and mesoscale eddy energy exchange is evaluated. Finally, theoretical solutions governing submesoscale front (SMF) – NIW interactions are developed and the validity of these solutions is tested against both idealized numerical simulations and realistic numerical simulation in the Gulf of Mexico. The main conclusions of this dissertation research are summarized as follows.

5.1 Conclusions

In the presence of a baroclinic geostrophic flow, the dispersion of NIWs is achieved mainly through the phase separation among different modes. Both eigen-frequency differences and mode-mode interferences contribute to the phase separation. In addition,

local interferences in modal space play a much more dominant role than non-local mode-mode interferences. A simplified version of the YBJ equation, i.e., the LLMI model, is constructed by dropping the non-local mode-mode interference terms and shows good skills in simulating the NIW dispersion in a background baroclinic geostrophic flow.

Analytical solutions derived from the LLMI model suggest that the eigen-frequency difference among different modes is reduced in the baroclinic case compared to the barotropic case, assuming the same surface values of geostrophic vorticity. But this reduction is largely compensated by the increase from local mode-mode interferences, leading to a comparable dephasing time scale to that in the barotropic case. As a result, the downward radiation of NIWs from the mixed layer to deep ocean is similar in the presence of baroclinic and barotropic flows with the same surface intensity.

Long-term mooring array deployed in the northern Gulf of Mexico during the GISR project is analyzed to gain understanding of energy exchange between internal waves and mesoscale eddies. In the subthermocline (245-450 m), the observed energy exchange between mesoscale eddies and internal waves is close to zero when the OW parameter is negative. However, there is a marked energy transfer of the rate of $O(10^{-10}m^2s^{-3})$ from mesoscale eddies to internal waves when the OW parameter is positive. NIWs play a dominant role in this energy exchange. This is due to both their energetic wave stress and strong interaction with mesoscale eddies.

This internal wave and mesoscale eddy energy exchange is well described by the wave capture mechanism. NIWs become highly anisotropic when the OW parameter is positive. The observed PDF of propagation direction of NIWs is consistent with the

prediction from the wave capture theory.

The simulated energy exchange between mesoscale eddies and NIWs in a high-resolution numerical simulation configured over the Gulf of Mexico is qualitatively consistent with the mooring observations. A permanent energy transfer from mesoscale eddies to NIWs is observed to mainly occurs when the OW parameter is positive, lending further supports to the dominant role of wave capture mechanism in the energy exchange between mesoscale eddies and NIWs.

Submesoscale O(<10 km) motions and their interactions with NIWs are studied theoretically and numerically using high-resolution ROMS simulations. Consistent with the recent modeling study by Zhong and Bracco (2013), SMFs with energetic vertical motions in ocean interior are found to be closely associated with NIWs. A dynamic mechanism explaining SMF development in the presence of background NIWs is presented. It shows that in convergence (downwelling) regions of NIWs, energy flux of the submesoscale motions converges and the energy is transferred from the NIWs to submesoscale motions, leading to enhanced submesoscale vertical velocity. The opposite is true in divergence (upwelling) zones of NIWs. The underlying dynamics can be understood in terms of wave action conservation of submesoscale internal gravity waves in the presence of background NIWs.

The validity of the theoretical finding is confirmed by high-resolution ROMS simulations forced by a realistic atmospheric forcing in the Gulf of Mexico. The simulations show that submesoscale vertical motions within convergence (downwelling) and divergence (upwelling) regions of NIWs exhibit a marked asymmetry with

significantly stronger submesoscale vertical velocity occurring in the convergence (downwelling) regions. By suppressing near-inertial wind forcing in the simulation, NIWs are substantially reduced, which further leads to a significant decrease in the submesoscale vertical motions, confirming that NIWs can contribute to the development of submesoscale vertical motions.

5.2 Future Work

Despite the progress made in this dissertation, there are still many questions remaining to be answered in future studies.

First, as documented in Chapter II, phase separations among NIW modes in the presence of barotropic and baroclinic geostrophic flows with the same surface intensity are comparable because the reduced eigen-frequency difference in the baroclinic flow is largely compensated by the mode-mode interferences. However, the underlying reasons for this compensation effect remain unknown and should be examined in future studies.

Second, analyses in Chapter III reveal that energy is irreversibility transferred from mesoscale eddies to NIWs through the wave capture mechanism, suggesting that NIWs act to damp mesoscale eddies. Future studies are required to quantify how important this damping effect is in energy budget of mesoscale eddies.

Third, as demonstrated in Chapter IV, intense submesoscale vertical motions can generated by NIWs. This suggests that NIWs may play an important role in the oceanic vertical transport processes. Further studies are needed to quantify the importance of this

vertical transport by directly simulating tracer transport processes using high-resolution models.

REFERENCES

Alford, M. H., and M. C. Gregg, 2001: Near-inertial mixing: modulation of shear, strain and microstructure at low latitude. Journal of Geophysical Research, 106, 16947–16968.

Alford, M. H., 2003a: Improved global maps and 54-years history of wind-work on ocean inertial motions. Geophysical Research Letters, 30, 1424.

Alford, M. H., 2003b: Redistribution of energy available for ocean mixing by long-range propagation of internal waves. Nature, 423, 159-162.

Alford, M. H. and M. Whitmont, 2007: Seasonal and spatial variability of near-inertial kinetic energy from historical moored velocity records. Journal of Physical Oceanography, 37, 2022–2037.

Alford, M. H., 2008: Observations of parametric subharmonic instability of the diurnal internal tide in the South China Sea. Geophysical Research Letters, 35, L15602.

Alford, M. H., M. F. Cronin, J. M. Klymak, 2012: Annual cycle and depth penetration of wind-generated near-inertial internal waves at ocean station Papa in the Northeast Pacific. Journal of Physical Oceanography, 42, 889–909.

Alford, M. H., A. Y. Shcherbina, and M. C. Gregg, 2013: Observations of near-inertial internal gravity waves radiating from a frontal jet. Journal of Physical Oceanography, 43, 1225-1239.

Alpers, W., 1985: Theory of radar imaging of internal waves. Nature, 314, 245-247.

Balmforth, N. J., S. G. Llewellyn Smith, and W. R. Young, 1998: Enhanced dispersion of near-inertial waves in an idealized geostrophic flow. Journal of Marine Research, 56, 1-40.

Balmforth, N. J., and W. R. Young, 1999: Radiative damping of near-inertial oscillations in the mixed layer. Journal of Marine Research, 57, 561-584.

Bretherton, F. P., 1966: The propagation of groups of internal gravity waves in a shear flow. Quarterly Journal of the Royal Meteorological Society, 92, 466–480.

Bretherton, F. P., and C. J. R. Garrett, 1968: Wavetrains in inhomogeneous moving media. Proceedings of the Royal Society, A302, 529-554.

Burwell, D., E. Tolkova, and A. Chawla, 2007: Diffusion and dispersion characterization of a numerical tsunami model. Ocean Modelling. 19, 10-30.

Bühler, O., and M. E. McIntyre, 2005: Wave capture and wave-vortex duality. Journal of Fluid Mechanics, 534, 67–95.

Brown, E. D., and W. B. Owens, 1981: Observations of the horizontal interactions between the internal wave field and the mesoscale flow. Journal of Physical Oceanography, 11, 1474–1480.

Chaigneau, A., O. Pizarro, and W. Rojas, 2008: Global climatology of near-inertial current characteristics from Lagrangian observations. Geophysical Research Letters, 35, L13603.

Chassignet, E. P., and Coauthors, 2007: The HYCOM (hybrid coordinate ocean model) data assimilative system. Journal of Marine Systems, 65, 60-83.

D'Asaro, E. A., 1985: The energy flux from the wind to near-inertial motions in the surface mixed layer. Journal of Physical Oceanography, 15, 1043–1059.

D'Asaro, E. A., 1989: The decay of wind-forced mixed layer inertial oscillations due to the β effect. Journal of Geophysical Research, 94, 2045–2056.

D'Asaro, E. A., and Coauthors, 1995: Upper-ocean inertial currents forced by a strong storm a strong storm. Part I: data and comparisons with linear theory. Journal of Physical Oceanography, 25, 2909–2936.

Danioux, E., J. Vanneste, P. Klein and H. Sasaki, 2012: Spontaneous inertia-gravity-wave generation by surface-intensified turbulence. Journal of Fluid Mechanics, 699, 153-173.

Dee, D. P. and Coauthors, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Quarterly Journal of the Royal Meteorological Society, 137, 553–597.

Egbert, G. D., and R. D. Ray, 2001: Estimates of M2 tidal energy dissipation from TOPEX/POSEIDON altimeter data. Journal of Geophysical Research, 106, 22475–22502.

Elipot, S. and R. Lumpkin, 2008: Spectral description of oceanic near-surface variability. Geophysical Research Letters, 35, L05606.

Ferrari, R., and C. Wunsch, 2009: Ocean circulation kinetic energy- reservoirs, sources, and sinks. Annual Review of Fluid Mechanics, 41, 253-282.

Fox-Kemper, B., R. Ferrari, and R. W. Hallberg, 2006: Modeling and parameterizing mixed layer eddies. Eos Trans. AGU 87 (36), Ocean Sci. Meet. Suppl., Abstract OS23F-06.

Frankignoul, C., 1976: Observed interaction between oceanic internal waves and mesoscale eddies. Deep-Sea Research, 23, 805–820.

Frankignoul, C., and T. M. Joyce, 1979: On the internal wave variability during the internal wave experiment (IWEX). Journal of Geophysical Research, 84, 769–776.

Fu, L.-L., 1981: Observations and models of inertial waves in the deep ocean. Review of Geophysics, 19, 141–170.

Furuichi, N., T. Hibiya, and Y. Niwa, 2008: Model-predicted distribution of wind-induced internal wave energy in the world's oceans. Journal of Geophysical Research, 113, C09034.

Garrett, C. and W. Munk, 1979: Internal waves in the ocean. Annual Review of Fluid Mechanics, 11, 339-369.

Garrett, C. J. R., and J. W. Loder, 1981: Dynamical aspects of shallow sea fronts. Philosophical Transactions of the Royal Society of London Series, A302, 563–581.

Garrett, C., 2001: What is the “near-inertial” band and why is it different from the rest of the internal wave spectrum? Journal of Physical Oceanography, 31, 962–971.

Gregory, J. M., 2000: Vertical heat transports in the ocean and their effect on time dependent climate change. Climate Dynamics, 16, 501-515.

Gill, A. E., 1984: On the behavior of internal waves in the wakes of storms. Journal of Physical Oceanography, 14, 1129-1151.

Greatbatch, R. J., 1984: On the response of the ocean to a travelling storm: parameters and scales. Journal of Physical Oceanography, 14, 59–78.

Hansen, C., and A. Samuelsen, 2009: Influence of horizontal model grid resolution on the simulated primary production in an embedded primary production model in the Norwegian Sea. Journal of Marine Systems, 75, 236–244.

Henyey, F. S., J. Wright, and S. M. Flatté, 1986: Energy and action flow through the internal wave field: An eikonal approach. Journal of Geophysical Research, 91, 8487–8495.

Hoskins, B. J., and F. P. Bretherton, 1972: Atmospheric frontogenesis models: mathematical formulation and solution. Journal of the Atmospheric Sciences, 29, 11–37.

Huang, R. X., 1999: Mixing and energetics of the oceanic thermohaline circulation. Journal of Physical Oceanography, 29, 727-746.

Jayne, S. R., and L. C. St. Laurent, 2001: Parameterizing tidal dissipation over rough topography. Geophysical Research Letters, 28, 811–814.

Jayne, S. R., 2009: The impact of abyssal mixing parameterizations in an ocean general circulation model. Journal of Physical Oceanography, 39, 1756-1775.

Jiang, J., Y. Lu, and W. Perrie, 2005: Estimating the energy flux from the wind to ocean inertial motions: the sensitivity to surface wind fields. Geophysical Research Letters, 32, L15610.

Jing, Z., and L. Wu., 2010: Seasonal variation of turbulent diapycnal mixing in the northwestern Pacific stirred by wind stress. Geophysical Research Letters, 37, L23604.

Jing, Z., and L. Wu, 2013: Low-frequency modulation of turbulent diapycnal mixing by anticyclonic eddies inferred from the hot time series. Journal of Physical Oceanography, 43, 824–835.

Jing Z., and L. Wu, 2014: Intensified diapycnal mixing in the midlatitude western boundary current. Scientific Reports, 4, 7412.

Jochum, M. and Coauthors, 2013: The impact of oceanic near-inertial waves on climate. Journal of Climate, 26, 2833–2844.

Johnson, R. S., 2005: Singular perturbation theory: mathematical and analytical techniques with applications to engineering. Springer, 292 pp.

Jones, W. L., 1969: Ray tracing for internal gravity waves. Journal of Geophysical Research, 74, 2028–2033.

Klein, P., and A. M. Treguier, 1995: Dispersion of wind-induced inertial waves by a barotropic jet. Journal of Marine Research, 53, 1-22.

Klein, P., and S. L. Smith, 2001: Horizontal dispersion of near-inertial oscillations in a turbulent mesoscale eddy field. Journal of Marine Research, 59, 697-723.

Klein, P., S. L. Smith, and G. Lapeyre, 2004: Organization of near-inertial energy by an eddy field. Quarterly Journal of the Royal Meteorological Society, 130, 1153-1166.

Kunze, E. 1985: Near-inertial propagation in geostrophic shear. Journal of Physical Oceanography, 15, 544–565.

Kunze, E., R.W. Schmitt, and J.M. Toole, 1995: The energy balance in a warm-core ring's near-inertial critical layer. Journal of Physical Oceanography, 25, 942-957.

Large, W. G., J. C. McWilliams, and S. C. Doney, 1994: Oceanic vertical mixing - a review and a model with a nonlocal boundary-layer parameterization. Review of Geophysics, 32, 363-403.

Lee, D.-K., and P. P. Niiler, 1998: The inertial chimney: the near-inertial energy drainage from the ocean surface to the deep layer. Journal of Geophysical Research, 103, 7579–7591.

Levy, M., P. Klein, A.-M. Treguier, 2001: Impacts of sub-mesoscale physics on production and subduction of phytoplankton in an oligotrophic regime. Journal of Marine Research, 59, 535–565.

Lighthill, J., 1978: Waves in fluids. Cambridge Univ. Press, 477 pp.

Locarnini, R. A., and Coauthors, 2010: World ocean atlas 2009, Volume 1: Temperature. S. Levitus, Ed. NOAA Atlas NESDIS 68, U.S. Government Printing Office, Washington, D.C., 184 pp.

Mahadevan, A., and D. Archer, 2000: Modeling the impact of fronts and mesoscale circulation on the nutrient supply and biogeochemistry of the upper ocean. Journal of Geophysical Research, 105, 1209–1225.

MacKinnon, J. A., and K. B. Winters, 2005: Subtropical catastrophe: significant loss of low-mode tidal energy at 28.9 °, Geophysical Research Letters, 32, L15605.

MacKinnon, J. A. and Coauthors, 2013: Parametric subharmonic instability of the internal tide at 29 °N. Journal of Physical Oceanography, 43, 17-28.

Martin, A. P., and P. Pondaven, 2003: On estimates for the vertical nitrate flux due to eddy-pumping. Journal of Geophysical Research, 108, 3359.

Molemaker, M. Jeroen, J. C. McWilliams, and I. Yavneh, 2005: Baroclinic instability and loss of balance. Journal of Physical Oceanography, 35, 1505–1517.

Munk, W., and C. Wunsch, 1998: Abyssal recipes II: energetics of tidal and wind mixing. Deep-Sea Research. I, 45, 1977-2010.

Müller, P., 1976: On the diffusion of momentum and mass by internal gravity waves. Journal of Fluid Mechanics, 77, 789–823.

Nagai, T., A. Tandon, E. Kunze, and A. Mahadevan, 2015: Spontaneous generation of near-inertial waves by the Kuroshio front. Journal of Physical Oceanography, 45, 2381–2406.

Nikurashin, M. and R. Ferrari, 2010a: Radiation and dissipation of internal waves generated by geostrophic motions impinging on small-scale topography: theory. Journal of Physical Oceanography, 40, 1055-1074.

Nikurashin, M. and R. Ferrari, 2010b: Radiation and dissipation of internal waves generated by geostrophic motions impinging on small-scale topography: application to the Southern Ocean. Journal of Physical Oceanography, 40, 2025-2042.

Olbers, D. J., 1976: Non-linear energy transfer and the energy balance of the internal wave field in the deep ocean. Journal of Fluid Mechanics, 74, 375-399.

Oschlies, A., 2002: Can eddies make ocean deserts bloom? Global Biogeochemical Cycles, 16, 1106.

Pedlosky, J., 1987: Geophysical fluid dynamics. Springer, 710 pp.

Plueddemann, A. J. and J. T. Farrar, 2006: Observations and models of the energy flux from the wind to mixed layer inertial currents. Deep-Sea Research. I, 53, 5-30.

Pollard, R. T., and R. C. Millard, 1970: Comparison between observed and simulated wind-generated inertial oscillations. Deep-Sea Research, 17, 153–175.

Polzin, K. L., 2010: Mesoscale eddy–internal wave coupling. Part ii: energetics and results from Polymode. Journal of Physical Oceanography, 40, 789–801.

Price, J. F., 1981: Upper ocean response to a hurricane. Journal of Physical Oceanography, 11, 153–175.

Price, J. F., 1983: Internal wave wake of a moving storm. Part I: scales, energy budget and observations. Journal of Physical Oceanography, 13, 949-965.

Price, J. F., R. A. Weller, and R. Pinkel, 1986: Diurnal cycling: observations and models of the upper ocean response to diurnal heating, cooling, and wind mixing. Journal of Geophysical Research, 91, 8411–8427.

Price, J. F., T. B. Sanford, and G. Z. Forristall, 1994: Forced stage response to a moving hurricane. Journal of Physical Oceanography, 24, 233–260.

Provenzale, A., 1999: Transport by coherent barotropic vortices. Annual Review of Fluid Mechanics, 31, 55–93.

Rainville, L. and R. Pinkel, 2004: Observations of energetic high-wavenumber internal waves in the Kuroshio. Journal of Physical Oceanography, 34, 1495–1505.

Rimac, A., J-S. von Storch, C. Eden, and H. Haak, 2013: The influence of high resolution wind stress field on the power input to near-inertial motions in the ocean. Geophysical Research Letters, 40, 4882–4886.

Richards, K. J., S.-P. Xie, T. Miyama, 2009: Vertical mixing in the ocean and its impact on the coupled ocean-atmosphere system in the eastern tropical Pacific. Journal of Climate, 22, 3703–3719.

Ruddick, B. R., and T. M. Joyce, 1979: Observations of interaction between the internal wave field and low-frequency flows in the North Atlantic. Journal of Physical Oceanography, 9, 498–517.

Saenko, O. and W. Merrifield, 2005: On the effect of topographically enhanced mixing on the global ocean circulation. Journal of Physical Oceanography, 35, 826-834.

Sarmiento, J. L. and J. R. Toggweiler, 1984: A new model for the role of the oceans in determining atmospheric pCO2. Nature, 308, 621-624.

Shchepetkin, A. F., and J. C. McWilliams, 2005: The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model. Ocean Modelling, 9, 347-404.

Shuto, N., 1991: Numerical simulation of tsunamis-its present and near future. Natural Hazzards, 4, 171-191.

Silverthorne, K. E., and J. M. Toole, 2009: Seasonal kinetic energy variability of near-inertial motions. Journal of Physical Oceanography, 39, 1035–1049.

Simmons, H. L. and M. H. Alford, 2012: Simulating the long-range swell of internal waves generated by ocean storms. Oceanography, 25, 30–41.

Stern, M. E., 1965: Interaction of a uniform wind stress with a geostrophic vortex. Deep Sea Research, 12, 355–367

Thomas, L. N., and C. M. Lee, 2005: Intensification of ocean fronts by down-front winds. Journal of Physical Oceanography, 35, 1086–1102.

van Meurs, P., 1998: Interactions between near-inertial mixed layer currents and the mesoscale: the importance of spatial variabilities in the vorticity field. Journal of Physical Oceanography, 28, 1363–1388.

Vanneste, J., 2013: Balance and spontaneous wave generation in geophysical flows. Annual Review of Fluid Mechnanics, 45, 147-172.

Vitousek, S., and O. B. Fringer, 2011: Physical vs. numerical dispersion in nonhydrostatic ocean modeling. Ocean Modelling, 40(1), 72–86.

Watanabe, M., and T. Hibiya, 2002: Global estimates of the wind-induced energy flux to inertial motions in the surface mixed layer. Geophysical Research Letters, 29, L1239.

Weller, R. A., 1982: The relation of near-inertial motions observed in the mixed layer during the JASIN (1978) experiment to the local wind stress and to the quasi-geostrophic flow field. Journal of Physical Oceanography, 12, 1122–1136.

Whalen, C. B., L. D. Talley, and J. A MacKinnon, 2012: Spatial and temporal variability of global ocean mixing inferred from Argo profiles. Geophysical Research Letters, 39, L18.

Whitt, D. B. and L. N. Thomas, 2015: Resonant generation and energetics of wind-forced near-inertial motions in a geostrophic flow. Journal of Physical Oceanography, 45, 181–208.

Wunsch, C., and R. Ferrari, 2004: Vertical mixing, energy and the general circulation of the oceans. Annual Review of Fluid Mechanics, 36, 281-412.

Young, W. R., and M. Ben Jelloul, 1997: Propagation of near-inertial oscillations through a geostrophic flow. Journal of Marine Research, 55, 735-766.

Zhai, X., R. J. Greatbatch, and J. Zhao, 2005: Enhanced vertical propagation of storm-induced near-inertial energy in an eddying ocean channel model. Geophysical Research Letters, 32, L18602.

Zhai, X., R. J. Greatbatch, and C. Eden, 2007: Spreading of near-inertial energy in a 1/12 ° model of the North Atlantic Ocean. Geophysical Research Letters, 34, L10609.

Zhai, X., R. J. Greatbatch, C. Eden, and T. Hibiya, 2009: On the loss of wind-induced near-inertial energy to turbulent mixing in the upper ocean. Journal of Physical Oceanography, 39, 3040–3045.

Zhang, X., S. F. DiMarco, D. C. Smith, IV, M. K., Howard, A. E. Jochens, and R. D. Hetland, 2009: Near-resonant ocean response to sea breeze on a stratified continental shelf. Journal of Physical Oceanography, 39, 2137-2155.

Zhang, X., D. C. Smith, IV, S. F. DiMarco, and R. D. Hetland, 2010: A numerical study of sea breeze driven ocean Poincare wave propagation and mixing near the critical latitude. Journal of Physical Oceanography, 40, 48-66.

Zhong, Y., and A. Bracco, 2013: Submesoscale impacts on horizontal and vertical transport in the Gulf of Mexico. Journal of Geophysical Research, 118, 5651–5668.

APPENDIX A

The YBJ equation is derived based on the multiple timescale approach (Johnson 1998). Its validity lies essentially in the slight departure of the wave frequency from the inertial frequency i.e., $\varepsilon^2 << 1$. Young and Ben Jelloul (1997) assumed that the geostrophic flow is characterized by a small Rossby number, a O(1) Burger number, and a vertical scale much larger than that of NIWs. In this case, the modulation of NIWs by geostrophic flows leads to a slight departure of the wave frequency from the inertial frequency. However, the assumptions can be further relaxed but still ensure $\varepsilon^2 << 1$ at the same time. In this appendix, we will demonstrate the validity of the YBJ equation under the assumption that the geostrophic flow is characterized by a small Rossby number, a O(1) Burger number, and an vertical/horizontal aspect ratio much larger than that of NIWs. Decompose motions into NIWs and geostrophic flow:

$$\mathbf{u}_t = \mathbf{u} + \mathbf{U} \tag{A1a}$$

$$b_t = b + B \tag{A1b}$$

$$p_t = p + P \tag{A1c}$$

where $\mathbf{u}_t$ is the three-dimensional velocity, b_t the buoyancy, and p_t the density-normalized pressure. Here and further on, the lowercase represents the NIWs with uppercase corresponding to the geostrophic flow. The linearized equations for NIWs are:

$$\frac{\partial u}{\partial t} + U\frac{\partial u}{\partial x} + V\frac{\partial u}{\partial y} + u\frac{\partial U}{\partial x} + v\frac{\partial U}{\partial y} + w\frac{\partial U}{\partial z} = (f_0 + \beta y)v - \frac{\partial p}{\partial x} \tag{A2a}$$

$$\frac{\partial v}{\partial t}+U\frac{\partial v}{\partial x}+V\frac{\partial v}{\partial y}+u\frac{\partial V}{\partial x}+v\frac{\partial V}{\partial y}+w\frac{\partial V}{\partial z}=-(f_0+\beta y)u-\frac{\partial p}{\partial y} \tag{A2b}$$

$$0=-\frac{\partial p}{\partial z}+b \tag{A2c}$$

$$\frac{\partial b}{\partial t}+U\frac{\partial b}{\partial x}+V\frac{\partial b}{\partial y}+u\frac{\partial B}{\partial x}+v\frac{\partial B}{\partial y}+w(\frac{\partial B}{\partial z}+N^2)=0 \tag{A2d}$$

To proceed systematically, it is essential to introduce non-dimensional variables. We choose L and H to represent the horizontal and vertical scales of the geostrophic motion while using l and h for NIWs. Let $u*$ and $U*$ denote the magnitudes of horizontal velocity of NIWs and geostrophic motions, respectively. N_0 represents the magnitude of background buoyancy frequency $N(z)$. Then we use these scalars to define non-dimensional dependent and independent variables, denoted by primes, as follows:

$$w=\frac{h}{l}u*w' \tag{A3a}$$

$$b=\frac{N_0^2}{f_0}\frac{h}{l}u*b' \tag{A3b}$$

$$p=\frac{N_0^2}{f_0}\frac{h^2}{l}u*p' \tag{A3c}$$

$$B=\frac{f_0 L}{H}U*B' \tag{A3d}$$

$$(x,y)=L(x',y') \tag{A3e}$$

$$z=Hz' \tag{A3f}$$

The non-dimensional equations are:

$$\frac{\partial u}{\partial t}+\delta_H R_0(U\frac{\partial u}{\partial \delta_H x}+V\frac{\partial u}{\partial \delta_H y})+R_0(u\frac{\partial U}{\partial x}+v\frac{\partial U}{\partial y})+R_0\frac{\delta_H}{\delta_V}w\frac{\partial U}{\partial z}= (1+R_0\hat{\beta}\cdot y)v-B_u\frac{\delta_H^2}{\delta_V^2}\frac{\partial p}{\partial \delta_H x} \tag{A4a}$$

$$\frac{\partial v}{\partial t}+\delta_H R_0(U\frac{\partial v}{\partial \delta_H x}+V\frac{\partial v}{\partial \delta_H y})+R_0(u\frac{\partial V}{\partial x}+v\frac{\partial V}{\partial y})+R_0\frac{\delta_H}{\delta_V}w\frac{\partial V}{\partial z}= -(1+R_0\hat{\beta}\cdot y)u-B_u\frac{\delta_H^2}{\delta_V^2}\frac{\partial p}{\partial \delta_H y} \tag{A4b}$$

$$0=-\frac{\partial p}{\partial \delta_V z}+b \tag{A4c}$$

$$\frac{\partial u}{\partial \delta_H x}+\frac{\partial v}{\partial \delta_H y}+\frac{\partial w}{\partial \delta_V z}=0 \tag{A4d}$$

$$\frac{\partial b}{\partial t}+\delta_H R_0(U\frac{\partial b}{\partial \delta_H x}+V\frac{\partial b}{\partial \delta_H y})+\frac{R_0}{B_u}\frac{\delta_V}{\delta_H}(u\frac{\partial B}{\partial x}+v\frac{\partial B}{\partial y})+w(\frac{R_0}{B_u}\frac{\partial B}{\partial z}+N^2)=0 \tag{A4e}$$

where $\delta_H=L/l$, $\delta_V=H/h$, $R_0=U^*/f_0L$, $B_u=N_0^2H^2/f_0^2L^2$, and $\hat{\beta}=\beta L^2/U^*$. For the sake of neatness, we use unprimed symbols for the non-dimensional variables here and hereinafter.

To ensure the validity of the YBJ equation, (A4) should satisfy the following relations to the first-order approximation:

$$\frac{\partial u}{\partial t}=v \tag{A5a}$$

$$\frac{\partial v}{\partial t}=-u \tag{A5b}$$

$$0=-\frac{\partial p}{\partial \delta_V z}+b \tag{A5c}$$

$$\frac{\partial u}{\partial \delta_H x}+\frac{\partial v}{\partial \delta_H y}+\frac{\partial w}{\partial \delta_V z}=0 \tag{A5d}$$

$$\frac{\partial b}{\partial t}+wN^2=0 \tag{A5e}$$

Young and Ben Jelloul (1997) assumed $R_0=\varepsilon^2<<1$, $B_u=O(1)$, $\delta_H=1$, $\delta_V=\varepsilon^{-1}$, and $\hat{\beta}\le O(1)$, in which case the O(1) terms in (A4) yield (A5). It can be easily demonstrated that the same conclusion also holds when $R_0=\varepsilon^2$, $B_u=O(1)$, $\delta_H=\varepsilon$, $\delta_V=1$, and $\hat{\beta}\le O(1)$. These two cases can be combined together, leading to a more general assumption i.e., $R_0=\varepsilon^2$, $B_u=O(1)$, $\delta_H/\delta_V=\varepsilon$, and $\hat{\beta}\le O(1)$. The condition $\delta_H/\delta_V=\varepsilon$ requires that the geostrophic flow is characterized by a horizontal/vertical aspect ratio much larger than that of NIWs.

In our simulations, $L=(2\alpha)^{-1}=40$ km is close to the first baroclinic radius 30 km so that $B_u=O(1)$. Furthermore, the magnitude of geostrophic vorticity is always much smaller than f_0 in the experiment E1, E2, and E3 (Table 2.1), leading to a Rossby number much smaller than unity. At this stage, the validity of the YBJ equation is determined by δ_H/δ_V. As can be seen from the pressure gradient terms in (A4a) and (A4b), the smallness of δ_H/δ_V is both the necessary and sufficient condition for the wave frequencies to be near-inertial.

δ_H (δ_V) increases with the horizontal (vertical) mode number *m* (*n*). Therefore, the smallness of δ_H/δ_V only holds for specific combinations of *m* and *n*. As revealed by

Figure 2.8 and 2.9, most of the initial near-inertial energy is projected onto the modes characterized by $|\omega_{m,n}| << f_0$ or equivalently $\delta_H / \delta_V << 1$. Furthermore, as local model-model interferences dominate the nonlocal mode-mode interferences, modes violating the smallness of δ_H / δ_V are unlikely to be excited within a short time, e.g., 10 IPs. The validity of the YBJ equation can be further evaluated by dropping the modes with $\omega_{m,n}$ comparable to f_0. We re-run the LLMI model by dropping the modes with $|\omega_{m,n}| > 0.2 f_0$. There is no significant difference compared to the original computation with all the modes included (Figure A1).

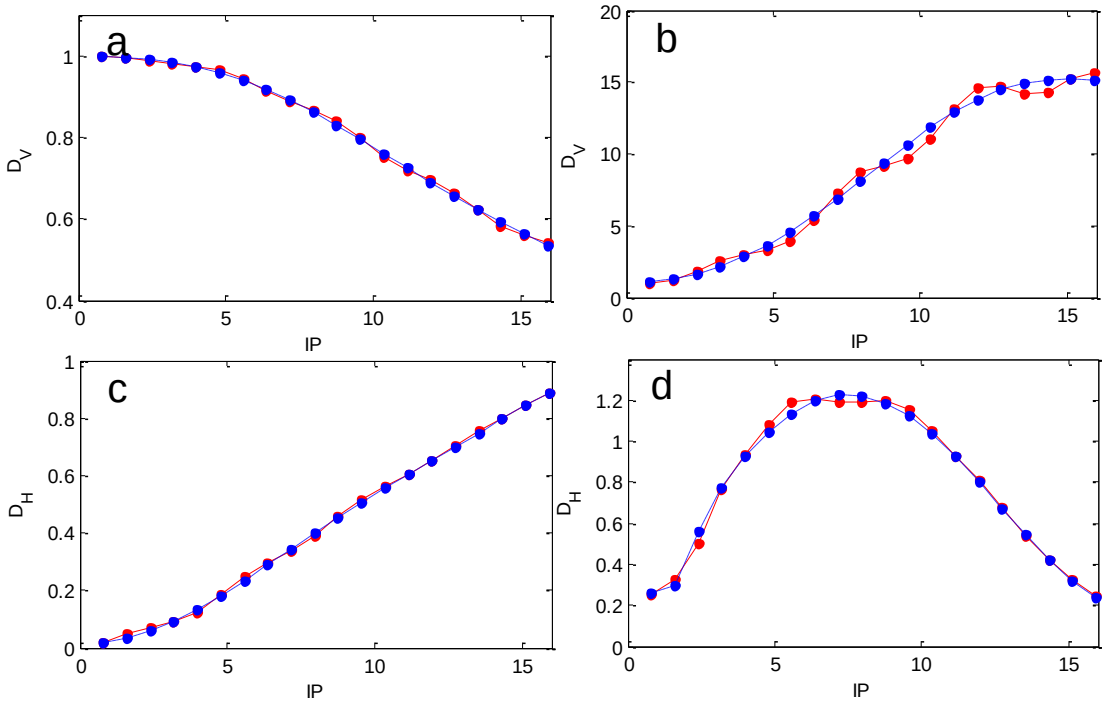

Figure A1 Dispersion of NIWs simulated by the LLMI model by using all the modes (red solid) and only the modes with $|\omega_{m,n}| < 0.2 f$ (blue dashed). (a) and (c) display the dispersion of NIWs in the mixed layer while (b) and (d) are in the deep ocean.

APPENDIX B

Under the geostrophic approximation, (3.1) and (3.2) can be written in the matrix form:

$$\frac{d}{dt}\begin{Bmatrix} k \\ l \end{Bmatrix} = -\frac{1}{2}\begin{bmatrix} S_n & S_s \\ S_s & -S_n \end{bmatrix}\begin{Bmatrix} k \\ l \end{Bmatrix} - \frac{1}{2}\begin{bmatrix} 0 & \varsigma \\ -\varsigma & 0 \end{bmatrix}\begin{Bmatrix} k \\ l \end{Bmatrix} \tag{B1}$$

where $S_n = U_x - V_y$ is the normal component of strain, $S_s = U_y + V_x$ the shear component, and $\varsigma = V_x - U_y$ the relative vorticity. Here S_n, S_s, and ς are assumed to be constant along the ray of NIWs. As the strain matrix is symmetric, it is always possible to choose the directions of the orthogonal axes (i.e., the principal axes) of reference so that the non-diagonal elements are zero:

$$\frac{d}{dt}\begin{Bmatrix} k', \\ l \end{Bmatrix} = -\frac{1}{2}\begin{bmatrix} S_n' & 0 \\ 0 & -S_n' \end{bmatrix}\begin{Bmatrix} k' \\ l' \end{Bmatrix} - \frac{1}{2}\begin{bmatrix} 0 & \varsigma' \\ -\varsigma' & 0 \end{bmatrix}\begin{Bmatrix} k' \\ l' \end{Bmatrix} \tag{B2}$$

where variables in the principle axes are denoted by primes. Note that the total strain variance and relative vorticity are invariants under the rotational transform so that $S_n'^2 = S_n^2 + S_s^2$ and $\varsigma' = \varsigma$.

Re-express the horizontal wave vector in polar coordinates:

$$(k', l') = k_h(\cos\theta', \sin\theta') \tag{B3}$$

where k_h is the magnitude of horizontal wavenumber and θ' is azimuth. Substituting (B3) into (B2) yields

$$\frac{dk_h}{dt} = -\frac{S_n'}{2} k_h \cos 2\theta' \tag{B4}$$

$$\frac{d\theta'}{dt} = \frac{S_n'}{2}\sin 2\theta' + \frac{\varsigma'}{2} \tag{B5}$$

(B5) indicates that both the strain and relative vorticity contribute to the rotation of wave vector. While the vorticity-induced angular velocity is constant (Figure B1a), the strain-induced angular velocity depends on θ'. The strain always tends to rotate the wave vector towards the direction along which the background flow converges (Figure B1b). This is because the phase contours in the direction where the background flow converge (diverge) are squeezed (stretched), leading to increased (decreased) magnitude of wave vector component in that direction.

When the OW parameter is negative ($S_n'^2 < \varsigma'^2$), the strain-induced angular velocity is always smaller than the vorticity-induced angular velocity (Figure B1c). In this case, the wave vector rotates unceasingly and its magnitude oscillates. However, strain- and vorticity-induced angular velocity are able to cancel each other when the OW parameter is positive ($S_n'^2 > \varsigma'^2$) (Figure B1d). There are four fixed points for θ' in this case, among which two are attractors ($\theta' = \theta_a'$ and $\theta' = \theta_a' + \pi$) and two are repellors ($\theta' = \theta_r'$ and $\theta' = \theta_r' + \pi$). It can be demonstrated that k_h will increase exponentially with time at attractors but decrease exponentially at repellors. Therefore, for almost all initial conditions, θ' asymptotically approaches either of its attractors and k_h increases exponentially with time, leading to the wave capture.

In the principal axes, the energy transfer rate (3.6) is reduced to

$$\varepsilon = -\int P_{u'u'} - P_{v'v'} d\omega \frac{S_n'}{2} \tag{B6}$$

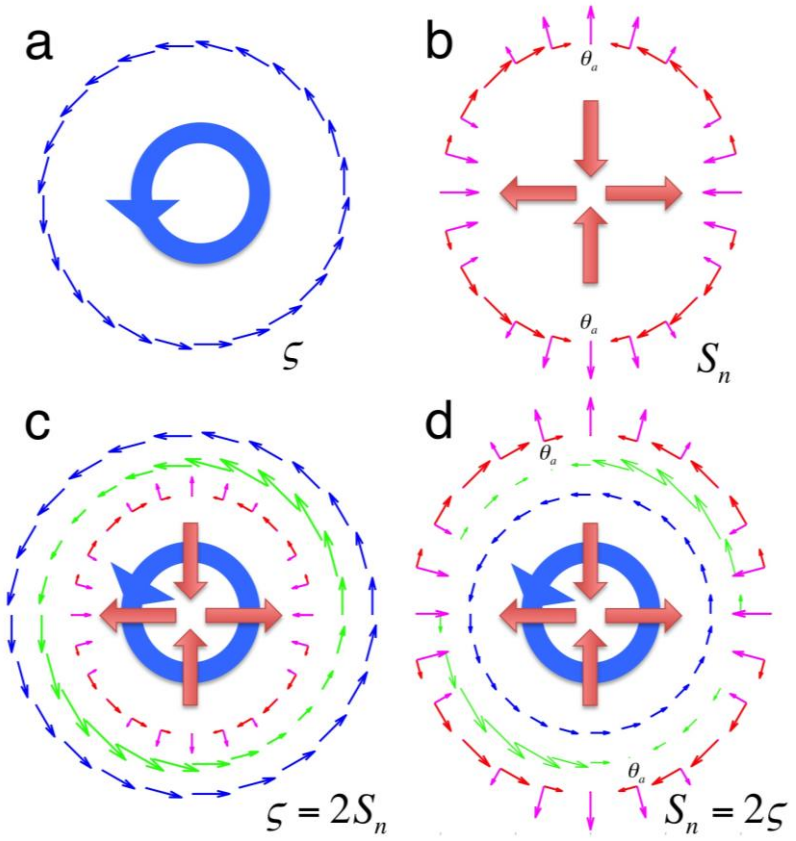

Figure B1 A schematic illustrating the influence of strain and vorticity of background flow on horizontal wave vector of internal waves. The blue and red thick arrows illustrate the vorticity and strain of background flow. The tangential thin arrows represent the temporal change of azimuth of horizontal wave vector. The radial thin arrows represent the change of wavenumber magnitude (pointing outwards means increase). (a), (b), (c), and (d) correspond to the case with vorticity only, strain only, vorticity stronger than strain, and strain stronger than vorticity, respectively.

Substituting (3.7) into (B6) yields

$$\varepsilon = -\int \frac{1-\tan^2\theta'}{1+\tan^2\theta'} \frac{\omega^2 - f^2}{\omega^2 + f^2} (P_{u'u'} + P_{v'v'}) d\omega \frac{S_n'}{2} \tag{B7}$$

In the wave capture scenario, $\sin 2\theta' = -S_n'/\varsigma'$ and $\cos 2\theta' < 0$. It follows:

$$\tan^2\theta' = \frac{(S_n' + \sqrt{S_n'^2 - \varsigma'^2})^2}{\varsigma'^2} \tag{B8}$$

It can be easily demonstrated that $\tan^2\theta' < 1$ when $S_n' < 0$ and $\tan^2\theta' > 1$ when $S_n' > 0$. Therefore, ε is always positive in the wave capture scenario, suggesting that energy of mesoscale eddies is transferred to internal wave field. The underlying dynamics can be understood in terms of wave action conservation of internal waves in the presence of background flow. In the wave capture scenario, the magnitude of horizontal wavenumber is amplified, contributing to increased intrinsic wave frequency as indicated by the dispersion relation of internal waves. In this case, the wave energy is required to increase accordingly to make the wave action conserved. This leads to a permanent energy transfer from mesoscale eddies to internal waves.

Finally, it should be noted that there are two attractors θ_a' and $\theta_a' + \pi$ for θ' in the wave capture scenario. But these two attractors lead to the identical ε as ε only depends on $\tan^2\theta'$. In the sense of energy transfer, we do not need to distinguish these two attractors.

APPENDIX C

In this appendix, we generalize the results in Section 3.1 to the two dimensional case in which the *y*-derivatives are not zero. The non-dimensional equations are:

$$\frac{\partial u}{\partial t}+\delta(U\frac{\partial u}{\partial x}+V\frac{\partial u}{\partial y})=\varepsilon v-\frac{\partial \eta}{\partial x}-\varepsilon\delta(u\frac{\partial U}{\partial \varepsilon x}+v\frac{\partial U}{\partial \varepsilon y}) \tag{C1a}$$

$$\frac{\partial v}{\partial t}+\delta(U\frac{\partial v}{\partial x}+V\frac{\partial v}{\partial y})=-\varepsilon u-\frac{\partial \eta}{\partial y}-\varepsilon\delta(u\frac{\partial V}{\partial \varepsilon x}+v\frac{\partial V}{\partial \varepsilon y}) \tag{C1b}$$

$$\frac{\partial \eta}{\partial t}+\delta(U\frac{\partial \eta}{\partial x}+V\frac{\partial \eta}{\partial y})+(1+\delta Z)(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y})+\delta\varepsilon(u\frac{\partial Z}{\partial \varepsilon x}+v\frac{\partial Z}{\partial \varepsilon y})=-\delta\varepsilon\eta(\frac{\partial U}{\partial \varepsilon x}+\frac{\partial V}{\partial \varepsilon y}) \tag{C1c}$$

The non-dimensional energy equation associated with (C1) is

$$\frac{\partial e}{\partial t}+\delta\frac{\partial Ue}{\partial x}+\delta\frac{\partial Ve}{\partial y}=-\nabla\cdot\mathbf{F}+P_b+P_s \tag{C2}$$

where $e=(1+\delta Z)(u^2+v^2)/2+\eta^2/2$ is the energy density of SMFs, $\mathbf{F}=(1+\delta Z)\mathbf{u}\eta$ the energy flux of SMFs due to pressure work, and P_b (P_s) the potential (kinetic) energy exchange between NIWs and SMFs:

$$P_b=-\frac{1}{2}\delta\varepsilon\eta^2(\frac{\partial U}{\partial \varepsilon x}+\frac{\partial V}{\partial \varepsilon y})$$

$$P_s=-\delta\varepsilon(1+\delta Z)(uu\frac{\partial U}{\partial \varepsilon x}+uv\frac{\partial U}{\partial \varepsilon y}+uv\frac{\partial V}{\partial \varepsilon x}+vv\frac{\partial V}{\partial \varepsilon y})$$

Introduce the fast and slow variables:

$$\frac{\partial \varphi}{\partial t}=-\omega,\ \frac{\partial \varphi}{\partial x}=k,\ \frac{\partial \varphi}{\partial y}=l,T_1=\varepsilon t,\ T_2=\varepsilon^2 t,\ X=\varepsilon x,Y=\varepsilon y \tag{C3}$$

where ω and (k,l) can be treated as the local wave frequency and wavenumber of SMFs, respectively. The temporal and spatial derivatives can be expanded, in virtue of the chain rule, as:

$$\frac{\partial}{\partial t} = -\omega \frac{\partial}{\partial \varphi} + \varepsilon \frac{\partial}{\partial T_1} + \varepsilon^2 \frac{\partial}{\partial T_2} \tag{C4a}$$

$$\frac{\partial}{\partial x} = k \frac{\partial}{\partial \varphi} + \varepsilon \frac{\partial}{\partial X} \tag{C4b}$$

$$\frac{\partial}{\partial y} = l \frac{\partial}{\partial \varphi} + \varepsilon \frac{\partial}{\partial Y} \tag{C4c}$$

Expand the perturbed quantities in asymptotic series of ε. The O(1) terms of (C1) are:

$$-\omega_0 \frac{\partial u_0}{\partial \varphi} = -k \frac{\partial \eta_0}{\partial \varphi} \tag{C5a}$$

$$-\omega_0 \frac{\partial v_0}{\partial \varphi} = -l \frac{\partial \eta_0}{\partial \varphi} \tag{C5b}$$

$$-\omega_0 \frac{\partial \eta_0}{\partial \varphi} + k \frac{\partial u_0}{\partial \varphi} + l \frac{\partial v_0}{\partial \varphi} = 0 \tag{C5c}$$

The solutions to (C5) are

$$(u_0, v_0, \eta_0) = \mathrm{Re}\{A_0 e^{i\varphi} (\frac{k}{\omega_0}, \frac{l}{\omega_0}, 1)\} \tag{C6}$$

with the dispersion relation $\omega_0 = \sqrt{k^2 + l^2}$.

The O(ε) terms of (C1) are

$$-\omega_0 \frac{\partial u_1}{\partial \varphi} + k \frac{\partial \eta_1}{\partial \varphi} = -\frac{\partial u_0}{\partial T_1} + \omega_1 \frac{\partial u_0}{\partial \varphi} - \frac{\partial \eta_0}{\partial X} - \chi(kU + lV) \frac{\partial u_0}{\partial \varphi} + v_0 \tag{C7a}$$

$$-\omega_0 \frac{\partial v_1}{\partial \varphi} + l\frac{\partial \eta_1}{\partial \varphi} = -\frac{\partial v_0}{\partial T_1} + \omega_1 \frac{\partial v_0}{\partial \varphi} - \frac{\partial \eta_0}{\partial Y} - \chi(kU + lV)\frac{\partial v_0}{\partial \varphi} - u_0 \tag{C7b}$$

$$-\omega_0 \frac{\partial \eta_1}{\partial \varphi} + k\frac{\partial u_1}{\partial \varphi} + l\frac{\partial v_1}{\partial \varphi} = -\frac{\partial \eta_0}{\partial T_1} + \omega_1 \frac{\partial \eta_0}{\partial \varphi} - \frac{\partial u_0}{\partial X} - \frac{\partial v_0}{\partial Y} - \chi(kU + lV)\frac{\partial \eta_0}{\partial \varphi} - \chi Z(k\frac{\partial u_0}{\partial \varphi} + l\frac{\partial v_0}{\partial \varphi}) \tag{C7c}$$

To get rid of the singularity, the following equations must be satisfied:

$$2(\frac{\partial A_0}{\partial T_1} + c_{gx}^0 \frac{\partial A_0}{\partial X} + c_{gy}^0 \frac{\partial A_0}{\partial Y}) + A_0(\frac{\partial}{\partial X}c_{gx}^0 + \frac{\partial}{\partial Y}c_{gy}^0) = 0 \tag{C8a}$$

$$\omega_1 = \chi(kU + lV) + \frac{\chi}{2}\omega_0 Z \tag{C8b}$$

where $c_{gx}^0 = \partial\omega_0 / \partial k$ and $c_{gy}^0 = \partial\omega_0 / \partial l$ are the O(1) group velocity.

In the aid of (C8), the solutions to (C7) are:

$$u_1 = \mathrm{Re}\{-\frac{\chi}{2}Z\frac{k}{\omega_0}A_0 e^{i\varphi} - \frac{i}{\omega_0}(\frac{\partial}{\partial T_1}(A_0\frac{k}{\omega_0}) + \frac{\partial A_0}{\partial X} - A_0\frac{l}{\omega_0})e^{i\varphi}\} \tag{C9a}$$

$$v_1 = \mathrm{Re}\{-\frac{\chi}{2}Z\frac{l}{\omega_0}A_0 e^{i\varphi} - \frac{i}{\omega_0}(\frac{\partial}{\partial T_1}(A_0\frac{l}{\omega_0}) + \frac{\partial A_0}{\partial Y} + A_0\frac{k}{\omega_0})e^{i\varphi}\} \tag{C9b}$$

$$\eta_1 = 0 \tag{C9c}$$

Substituting (C6) and (C9) into (C2) and averaging over one wave period or wavelength yield:

$$\begin{aligned}\frac{\partial < e >}{\partial T_1} + \varepsilon\frac{\partial < e >}{\partial T_2} + \frac{\partial c_{gx} < e >}{\partial X} + \frac{\partial c_{gy} < e >}{\partial Y} = -\delta < e > (\frac{1}{2}(\frac{\partial U}{\partial X} + \frac{\partial V}{\partial Y}) \\ + \cos^2\alpha\frac{\partial U}{\partial X} + \sin\alpha\cos\alpha(\frac{\partial U}{\partial Y} + \frac{\partial V}{\partial X}) + \sin^2\alpha\frac{\partial V}{\partial Y}) + O(\varepsilon^2)\end{aligned} \tag{C10}$$

where $< e >= | A_0^2 | / 2 + O(\varepsilon^2)$ is the energy density averaged over one wavelength or wave period, $\alpha = \arctan(l/k)$ the azimuth of the horizontal wave vector, and the group velocity:

$$c_{gx} = \frac{\partial \omega_0 + \omega_1}{\partial k} = (1 + \frac{\delta}{2} Z)\cos\alpha + \delta U \tag{C11a}$$

$$c_{gy} = \frac{\partial \omega_0 + \omega_1}{\partial l} = (1 + \frac{\delta}{2} Z)\sin\alpha + \delta V \tag{C11b}$$

Introduce $\omega_i = (1 + \delta Z/2)\sqrt{k^2 + l^2}$ as the intrinsic wave frequency. According to the ray tracing relation, we have

$$\begin{aligned} \frac{1}{\omega_i}(\frac{\partial \omega_i}{\partial T_1} + \varepsilon \frac{\partial \omega_i}{\partial T_2} + c_{gx}\frac{\partial \omega_i}{\partial X} + c_{gy}\frac{\partial \omega_i}{\partial Y}) = -\delta(\frac{1}{2}(\frac{\partial U}{\partial X} + \frac{\partial V}{\partial Y}) \\ + \cos^2\alpha \frac{\partial U}{\partial X} + \sin\alpha\cos\alpha(\frac{\partial U}{\partial Y} + \frac{\partial V}{\partial X}) + \sin^2\alpha \frac{\partial V}{\partial Y}) + O(\varepsilon^2) \end{aligned} \tag{C12}$$

Substitute (C12) into (C10) yields

$$\frac{\partial < n >}{\partial T_1} + \varepsilon \frac{\partial < n >}{\partial T_2} + \frac{\partial c_{gx} < n >}{\partial X} + \frac{\partial c_{gy} < n >}{\partial Y} = O(\varepsilon^2) \approx 0 \tag{C13}$$

(C13) states that the wave action of SMFs is conserved in the presence of NIWs.

APPENDIX D

In this section, we demonstrate the validity of our analytical solution (4.16) based on the comparisons with numerical simulations using the reduced gravity model. Figure D1 shows the initial interface elevation superposed by rightward-propagating NIWs and localized submesoscale perturbations with small amplitude. The wave frequency of NIWs is set to be $1.1f$ with a horizontal current speed of 0.2 m s^{-1}. These values are well representative in the real ocean. Figure D2 shows evolution of $< e >$, $< n >$, c_g, and k computed from the numerical simulations and analytical solutions (4.15) and (4.16). The analytical solutions agree well with numerical simulations. The difference in $< e >$ and $< n >$ between the analytical solutions and numerical simulations is within 10% of their initial values. We conclude that the analytical solution is valid.

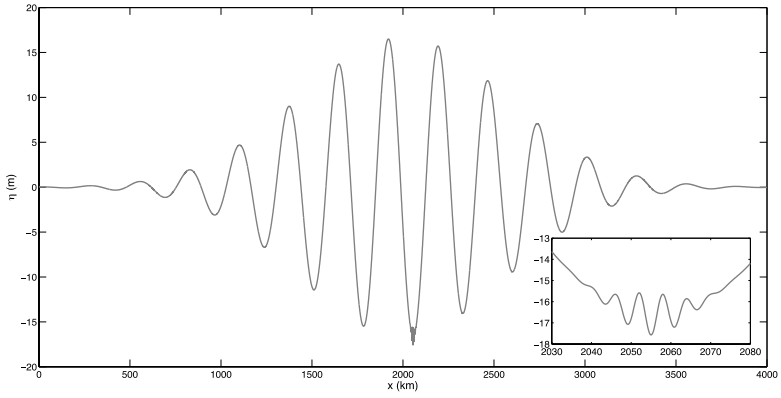

Figure D1 Initial condition for $Z + \eta$

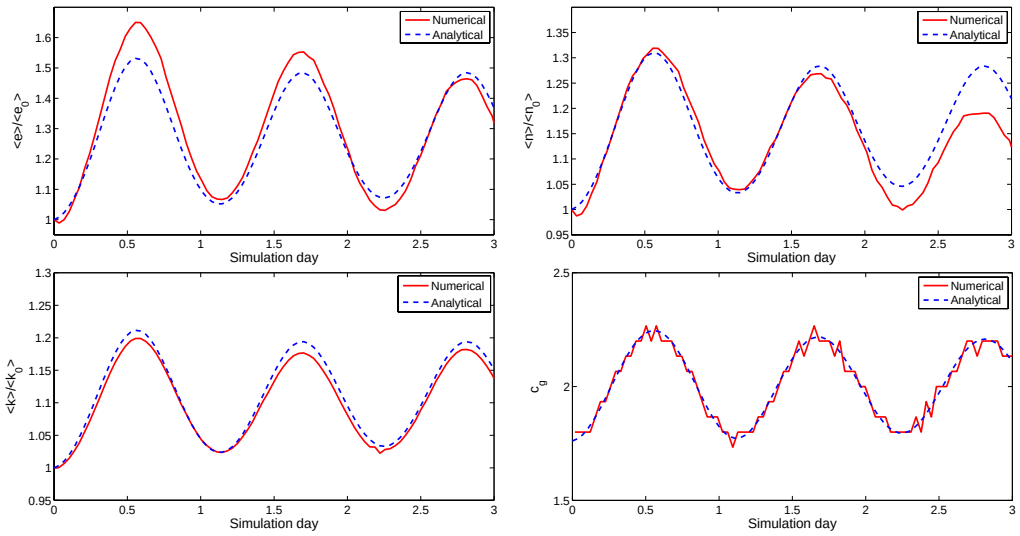

Figure D2 The evolution of energy density $<e>$ (upper left), action density $<n>$ (upper right), horizontal wavenumber k (lower left), and group velocity c_g (lower right). Here $<e>$, $<n>$, and k have been normalized by their initial values.

APPENDIX E

The O(1) terms of (4.21) are

$$-\omega_0 \frac{\partial u_1^0}{\partial \varphi} = -k \frac{\partial \eta_1^0}{\partial \varphi} \tag{E1a}$$

$$-\omega_0 \frac{\partial v_1^0}{\partial \varphi} = 0 \tag{E1b}$$

$$-\omega_0 \frac{\partial \eta_1^0}{\partial \varphi} + w_1^0 = 0 \tag{E1c}$$

$$k \frac{\partial u_1^0}{\partial \varphi} - w_1^0 = 0 \tag{E1d}$$

The solutions to (E1) are

$$(u_1^0, v_1^0, w_1^0, \eta_1^0) = \mathrm{Re}\{A_0 e^{i\varphi}(1,0,ik,1)\} \tag{E2}$$

with the associated dispersion relation: $\omega_0 = k$.

The $O(\varepsilon)$ terms of (4.21) are

$$-k\left(\frac{\partial u_1^1}{\partial \varphi} - \frac{\partial \eta_1^1}{\partial \varphi}\right) = i(\omega_1 - s\chi U_1 k)u_1^0 - \frac{\partial u_1^0}{\partial T_1} - \frac{\partial \eta_1^0}{\partial X} - \frac{s\chi}{2} w_1^0 U_1 \tag{E3a}$$

$$-k \frac{\partial v_1^1}{\partial \varphi} = -u_1^0 - \frac{s\chi}{2} w_1^0 V_1 \tag{E3b}$$

$$-k \frac{\partial \eta_1^1}{\partial \varphi} + w_1^1 = i\left(\omega_1 - \frac{s\chi}{2} U_1 k\right)\eta_1^0 - \frac{\partial \eta_1^0}{\partial T_1} - s\chi w_1^0 Z_1 \tag{E3c}$$

$$k \frac{\partial u_1^1}{\partial \varphi} - w_1^1 = -\frac{\partial u_1^0}{\partial X} \tag{E3d}$$

To get rid of the singularity, the following equation must be satisfied:

$$\frac{\partial A_0}{\partial T_1} + \frac{\partial c_g^0 A_0}{\partial X} = 0 \tag{E4a}$$

$$\omega_1 = s\chi k(U_1 + \frac{1}{2}Z_1) \tag{E4b}$$

where $c_g^0 = \partial\omega_0 / \partial k = 1$ is the O(1) group velocity of SMFs.

In the aid of (E4), the solutions to (E3) are

$$u_1^1 = \mathrm{Re}\{(-\frac{s\chi}{2}Z_1 + \frac{s\chi}{2}U_1)A_0 e^{i\varphi}\} \tag{E5a}$$

$$v_1^1 = \mathrm{Re}\{(-\frac{i}{k} + \frac{s\chi}{2}V_1)A_0 e^{i\varphi}\} \tag{E5b}$$

$$w_1^1 = \mathrm{Re}\{ik(-\frac{s\chi}{2}Z_1 + \frac{s\chi}{2}U_1)A_0 e^{i\varphi} + \frac{\partial A_0}{\partial X}e^{i\varphi}\} \tag{E5c}$$

$$\eta_1^1 = 0 \tag{E5d}$$

Substituting (E2) and (E5) into (4.22) and averaging over one wave period or wavelength yield:

$$< e_1 >= \frac{|A_0^2|}{2}(1 + \frac{s\delta}{2}U_1 - \frac{s\delta}{2}Z_1) + O(\varepsilon^2) \tag{E6a}$$

$$< P_{a,1} >= -s\delta\varepsilon\frac{3}{4}\frac{\partial U_1 < e_1 >}{\partial X} + O(\varepsilon^3) \tag{E6b}$$

$$< P_{w,1} >= -\varepsilon\frac{\partial}{\partial X} < e_1 > + O(\varepsilon^3) \tag{E6c}$$

$$< P_{s,1}^H >= -s\delta\varepsilon < e_1 > \frac{\partial U_1}{\partial X} + O(\varepsilon^3) \tag{E6d}$$

$$< P_{s,1}^V >= -s\delta\varepsilon(\frac{1}{4}U_1\frac{\partial}{\partial X} < e_1 > -\frac{1}{2} < e_1 > V_1) + O(\varepsilon^3) \tag{E6e}$$

$$< P_{b,1} >= -s\delta\varepsilon((\frac{1}{2}\frac{\partial Z_1}{\partial X} + \frac{3}{4}\frac{\partial U_1}{\partial X}) < e_1 > + \frac{1}{2} Z_1 \frac{\partial < e_1 >}{\partial X}) + O(\varepsilon^3) \quad \text{(E6f)}$$

Substituting (E6) into (4.22) yields (4.23) and (4.24).

Introduce the intrinsic wave frequency $\omega_i = k(1 + s\delta Z_1 / 2)$. According to the ray tracing relation (Lighthill 1978), we have

$$\frac{1}{\omega_i}(\frac{\partial \omega_i}{\partial T_1} + \varepsilon \frac{\partial \omega_i}{\partial T_2} + c_g \frac{\partial \omega_i}{\partial X}) = -\delta \frac{3}{2}\frac{\partial U}{\partial X} + O(\varepsilon^2) \quad \text{(E7)}$$

The equation for wave action density $n_1 = e_1 / \omega_i$ is:

$$\frac{\partial < n_1 >}{\partial T_1} + \varepsilon \frac{\partial < n_1 >}{\partial T_2} + \frac{\partial c_g < n_1 >}{\partial X} = \delta < n_1 > \frac{s}{2} V_1 + O(\varepsilon^2) \quad \text{(E8)}$$